sustainability

MDPI

Article

Can Public Data Openness Promote Inclusive Green Growth in Chinese Cities—Empirical Evidence Based on Double Machine Learning

Zongyi Ying [1], Weiye Liang [2,*], Linsen Zhu [2,*] and Zuowen Liao [2]

[1] Beibu Gulf Ocean Development Research Center, Beibu Gulf University, Qinzhou 535011, China
[2] The School of Economics and Management, Beibu Gulf University, Qinzhou 535011, China
* Correspondence: 18377788420@163.com (W.L.); zls@bbgu.edu.cn (L.Z.)

Abstract

Against the backdrop of the deepening digital economy and accelerated green and low-carbon transformation, data factors have emerged as a core strategic resource driving the formation and allocation of new quality productive forces. Whether data can transition from closure to openness, thereby converting data dividends into inclusive green growth momentum that balances efficiency and equity, constitutes a strategic imperative in the process of Chinese-style modernization. Based on city-level data in China from 2010 to 2024, this paper measures the inclusive green growth (IGG) levels of 295 prefecture-level and above cities, and employs the double/debiased machine learning (DDML) model to examine the impact and underlying mechanisms of public data open platforms on urban IGG, treating the launch of these platforms as a quasi-natural experiment. The findings reveal that the launch of public data open platforms exerts a significant positive effect on urban IGG and is associated with improvements across multiple dimensions, including economic development, social equity, green production and consumption, and environmental protection. In terms of mechanisms, public data open platforms promote IGG through three pathways: advancing digital inclusive finance development, mitigating factor market distortions, and improving urban information circulation. The impact effects exhibit significant heterogeneity across cities with different endowments: cities with larger economic scales, more abundant human capital, and higher initial IGG levels benefit more from public data openness. Furthermore, the effect of public data open platforms on urban IGG displays a regional gradient pattern of "strongest in the east, moderate in the west, and weakest in the central region," indicating that policy dividends are fully realized in information-intensive economies but remain insufficiently realized in regions with relatively lagging digital infrastructure.

Keywords: public data open platform; inclusive green growth; double machine learning; digital inclusive finance; factor market distortion; information circulation

Check for updates

Academic Editor: Weixin Yang

Received: 26 August 2026
Revised: 15 September 2026
Accepted: 23 September 2026
Published: 26 September 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) license.

1. Introduction

The report of the 20th National Congress of the Communist Party of China (CPC) explicitly states that "from this point forward, the central task of the Party is to unite and lead the people of all ethnic groups nationwide in comprehensively building a great modern socialist country and achieving the Second Centenary Goal, advancing the great rejuvenation of the Chinese nation through Chinese-style modernization" [1]. Currently, China's economy is at a critical juncture of transition from factor-driven to innovation-driven growth

and from rapid expansion to high-quality development. The deep integration of market-oriented allocation of data factors with green development is increasingly becoming an important force reshaping regional economic patterns and narrowing regional disparities. According to the National Bureau of Statistics of China, the value added of core industries of the digital economy as a share of GDP increased from 6.8% in 2015 to 10.5% in 2024, a cumulative increase of 3.7 percentage points over the decade. However, compared with the inherent requirements of high-quality development, the vigorous development of the digital economy still faces structural contradictions: on the one hand, the potential value of data factors is far from being fully released; on the other hand, green transformation and inclusive growth face dual pressures. The "Outline of the Fifteenth Five-Year Plan for National Economic and Social Development of the People's Republic of China" proposes to "accelerate the comprehensive green transformation of economic and social development and build a Beautiful China," that is, to pursue coordinated economic growth and environmental protection with development outcomes shared by all people. Therefore, how to achieve inclusive green growth while promoting both effective qualitative improvement and reasonable quantitative growth of the economy is an unavoidable issue in China's journey toward socialist modernization and common prosperity for all people.

IGG refers to a synergistic development model that simultaneously achieves economic growth, equal social opportunities, and environmental sustainability [2,3], integrating two core concepts: inclusive growth and green growth. The former emphasizes that the benefits of growth should reach all social groups and narrow income and opportunity gaps; the latter requires reducing carbon emissions and resource consumption while improving production efficiency. In the Chinese policy context, IGG is highly consistent with "common prosperity" and "modernization of harmony between humanity and nature," serving as an important benchmark for measuring urban high-quality development. Public data open platforms are portal vehicles through which local governments make non-sensitive public datasets available to the public free of charge in machine-readable formats, covering data on economic statistics, environmental monitoring, corporate credit, and public services. As shown in Figure 1, the launch of public data open platforms in China from 2012 to 2024 exhibits significant spatiotemporal diffusion characteristics: in the early stage, scattered pilot programs were concentrated in Beijing, Shanghai, and Guangdong; after 2017, the expansion accelerated, with Shandong, Guangdong, Guangxi, and Sichuan launching platforms in succession; the period from 2018 to 2019 witnessed a peak, followed by gradual extension to central, western, and northeastern regions. This differentiated progression across provinces and over time provides a natural quasi-experimental setting for identifying policy effects using a difference-in-differences framework.

Based on this, this paper takes 295 prefecture-level and above cities in China as the research subject. Using city-level panel data from 2010 to 2024, the entropy method is employed to comprehensively measure the urban IGG index. Treating the launch timing of local public data open platforms as a quasi-natural experiment, the DDML model is applied to examine the impact of public data open platforms on urban IGG. The internal transmission mechanisms are further revealed through three pathways: digital inclusive finance development, factor market distortion, and information circulation level and the heterogeneity characteristics of urban resource endowments and regional locations are systematically examined. The potential marginal contributions of this paper are threefold. First, by incorporating public data open platforms into the IGG analytical framework, this paper elucidates the promoting effect of public data openness on urban IGG from both theoretical and empirical perspectives, thereby extending the research boundary of IGG determinants and providing a new dimension for evaluating the comprehensive benefits of data factor market-oriented reform. Second, the DDML model is employed for causal

identification of the policy effects of public data openness, effectively mitigating the "curse of dimensionality" and model specification bias issues faced by traditional econometric models, thereby enhancing the credibility of the research findings. Third, the internal mechanisms through which public data open platforms affect IGG are systematically revealed from three dimensions—factor market distortion, digital inclusive finance development, and information circulation level and the heterogeneity characteristics of urban economic scale, human capital level, and regional location are further examined, providing empirical evidence for advancing public data open platform construction in a locally adapted manner. The remainder of this paper is organized as follows. Section 2 reviews the literature and presents the theoretical analysis. Section 3 describes the policy background and research hypotheses. Section 4 presents the research design, including the DDML model specification, variable selection, and data sources. Section 5 reports and discusses the baseline regression results and robustness checks. Section 6 provides an in-depth analysis of the mechanisms and heterogeneity characteristics. Section 7 offers further discussion. Section 8: Discussion. Section 9 concludes with research conclusions and policy recommendations.

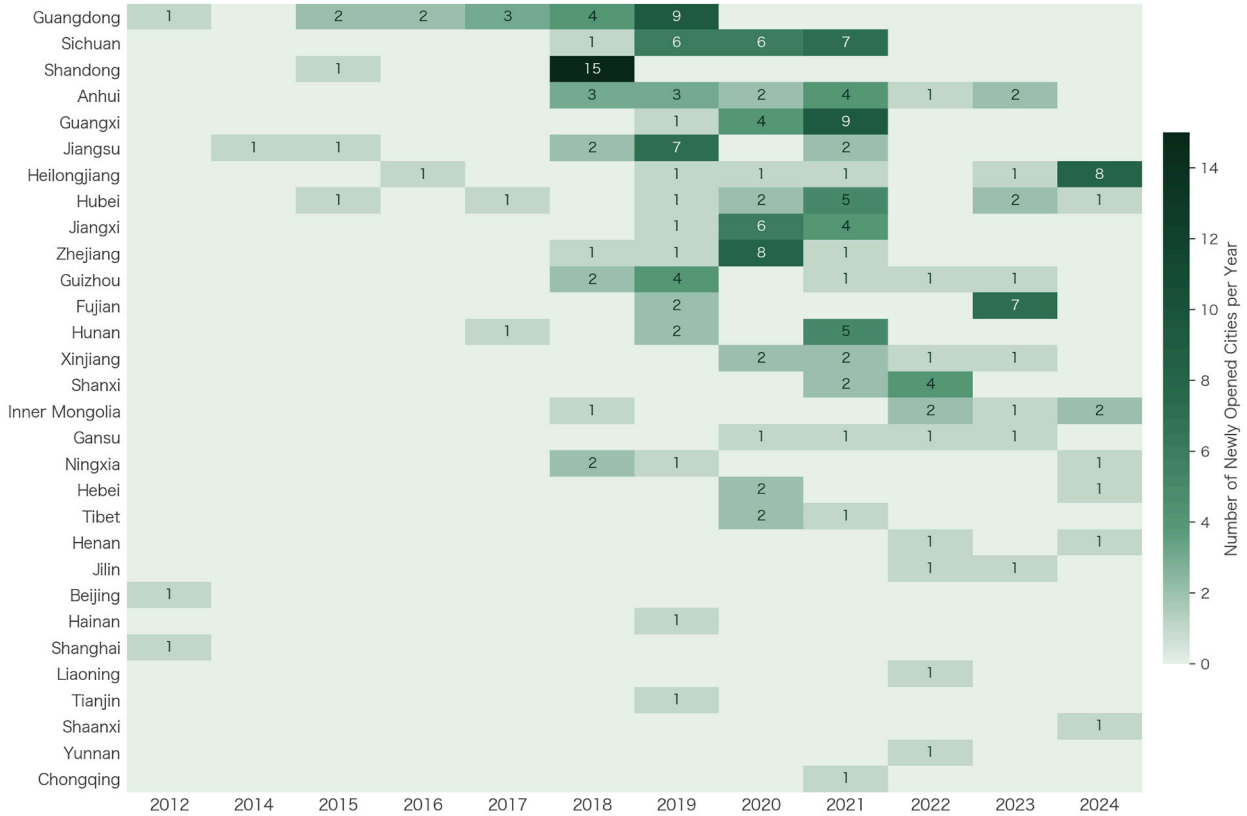


Figure 1. Spatiotemporal distribution of public data open platforms across Chinese provinces, 2012–2024. Note: The underlying sample period is 2010–2024. However, no prefecture-level city launched a public data open platform during 2010–2011, so the number of new platforms equals zero for all provinces in those two years. The heatmap therefore displays the spatiotemporal distribution from 2012 to 2024. Each cell reports the number of newly launched platforms in a given province-year.

2. Literature Review

2.1. Determinants of Inclusive Green Growth

IGG emphasizes the synergistic advancement of economic growth, social equity, and environmental sustainability, serving as an important composite indicator for measuring the level of high-quality development. Existing research has primarily focused on digital infrastructure, data factors, green innovation, financial development, and low-carbon policies.

First, digital infrastructure is considered an important foundational condition for promoting IGG. Network infrastructure construction can reduce information acquisition costs, improve resource allocation efficiency, and enhance urban green governance capacity by promoting digital technology diffusion. Related studies have shown that network infrastructure construction has a significant promoting effect on IGG [4,5]. Second, as a new type of production factor in the digital economy era, data factors possess characteristics such as non-rivalry, low reproduction cost, economies of scale, and externalities, and can influence urban high-quality development through improving production efficiency, promoting information sharing, and optimizing resource allocation [6,7]. At the city level, data factor agglomeration can not only promote green technology innovation and production efficiency improvement but also alleviate labor factor misallocation by facilitating the flow of high-quality talent and enhancing workers' digital literacy, thereby improving IGG levels [8,9]. Furthermore, the development of digital inclusive finance has been confirmed to promote inclusive green growth through channels such as alleviating financing constraints, expanding financial service availability, and supporting green innovation [10,11].

International studies have also provided empirical evidence from perspectives including policy pilots, digital infrastructure, and data factor marketization. Ai et al. [12], based on a double machine learning framework, found that low-carbon city pilot policies significantly improved urban IGG levels. Dong et al. [10] further pointed out that synergistic effects exist between digital infrastructure and energy transition policies, which can jointly drive inclusive green growth. Similarly, Shen et al. [13] and Dong et al. [14], approaching from the perspectives of data factor marketization and urban green growth, respectively, found that data factor market construction can improve green development performance through pathways such as enhancing carbon emission efficiency, promoting green technology innovation, and developing digital inclusive finance.

2.2. Economic Effects of Public Data Open Platforms

Public data open platforms are important institutional arrangements through which governments release public data resources and promote the social utilization of data factors. By reducing information acquisition costs, enhancing data availability, and strengthening information transparency, public data openness can exert multidimensional impacts on urban innovation, economic resilience, green development, and social equity.

International research indicates that the value created by open government data depends not only on data release itself, but also on data quality, accessibility, interoperability, institutional arrangements, and users' capacity to transform data into public and economic value [15,16]. European scholarship further conceptualizes open data as part of a sustainable urban data ecosystem rather than an isolated portal. Evidence from 19 European cities shows that the sustainability of open-data platforms depends on interactions among data providers, users, governance arrangements, and urban innovation actors [17]. From a sustainability-governance perspective, accessible environmental data can support evidence-based planning, environmental accountability, and decarbonization objectives; however, these benefits require appropriate legal and governance safeguards [18].

In terms of innovation effects, public data openness helps stimulate urban innovation and entrepreneurial vitality. On the one hand, open data can mitigate information asymmetry among governments, enterprises, and market entities, reducing institutional transaction costs. On the other hand, the availability of public data resources enhances enterprises' ability to identify market demand, obtain financing support, and conduct green technology innovation. Prior research consistently indicates that greater public data openness is associated with enhanced urban green innovation, increased entrepreneurial dynamism, and stronger firms' innovation capacity [19–24]. International research has also

found that government data openness and digital infrastructure construction have positive effects on firm entry, urban innovation, and financing constraint alleviation [25–28]. The release of innovation vitality further strengthens cities' ability to withstand external shocks. By improving the information transparency of urban governance and the efficiency of supply-demand matching, public data openness enables governments and market entities to more swiftly identify risks, allocate resources, and respond to external shocks. Related studies have found that public data openness enhances urban economic resilience through mechanisms such as collaborative governance, supply-demand adaptation, and supply chain optimization, with certain spatial spillover effects [28,29].

In the green development dimension, data factors can be embedded throughout the production, circulation, and governance processes, reducing the information costs of green innovation and environmental governance, and improving the precision of pollution identification, energy management, and pollution and carbon reduction. Studies have shown that public data openness can promote the coordinated development of urban digitalization and greenization, and drive green transformation by stimulating green innovation [19,30–32]. Related research on the digital economy and data factor marketization has also confirmed that the optimization of data factor allocation can generate energy-saving, efficiency-enhancing, and green growth effects [33–35]. Moreover, public data openness has the potential to promote social equity. By improving public service provision, reducing information barriers, and facilitating urban–rural and interregional factor flows, public data openness helps enhance the inclusiveness of resource allocation, thereby promoting common prosperity, urban–rural integration, and coordinated regional development [36].

2.3. Research Gap

Synthesizing the above literature, while existing studies have provided an important foundation for understanding the economic effects of public data openness, several gaps remain. First, few studies have directly incorporated public data open platforms and IGG into a unified analytical framework. Existing research has examined the effects of public data openness from single dimensions such as green innovation [19], economic resilience, and common prosperity [36], whereas IGG, as a composite indicator integrating economic growth, social equity, and environmental sustainability, can more comprehensively evaluate the policy dividends of public data openness. Second, most existing studies characterize policy benefits using single-dimensional indicators such as green innovation, economic resilience, and common prosperity, without evaluating the overall policy dividends of public data open platforms from a more multidimensional composite perspective. While single dimensions can reveal policy effects in specific domains, they struggle to capture the synergistic contributions of data openness to the "growth-inclusiveness-greenness" triad, leading to fragmented assessments. Third, in terms of identification strategies, most relevant studies employ the traditional difference-in-differences method, which faces potential model misspecification and high-dimensional control variable selection bias issues. Few studies have introduced more advanced causal inference methods such as double machine learning into the policy evaluation of public data open platforms, leaving the credibility of existing estimation results to be further improved. The DDML method can effectively mitigate the "curse of dimensionality" and functional form assumption bias through cross-fitting and flexible machine learning algorithms [37,38], but it has not yet been applied to the policy evaluation of public data openness.

3. Policy Background and Research Hypotheses

3.1. Policy Background

As a new type of production factor, data has become a strategic resource for promoting high-quality economic and social development. Since the early 21st century, public data openness has gradually evolved from an academic advocacy to an important governance agenda for governments worldwide. At the international level, the United States launched Data.gov in 2009, and the United Kingdom introduced data.gov.uk in 2010, from which the open government data movement rapidly spread globally. A systematic review by Safarov et al. [39] noted that open government data has produced positive effects in multiple dimensions, including innovation incentives, public service optimization, and economic growth. Hughes-Cromwick and Coronado [40] further confirmed the important value of government data for business decision-making from an enterprise decision perspective. At the domestic level, Shanghai pioneered the launch of a public data open platform in China in 2012. According to statistics from the Fudan University Laboratory for Digital and Mobile Governance, as of July 2025, a total of 257 local governments nationwide had launched public data open platforms, with the total number of open datasets growing nearly 57-fold compared to 2017 and total data volume exceeding 190.7 billion entries [41], indicating that public data openness has transitioned from "pilot exploration" to "comprehensive rollout."

In terms of top-level policy design, the CPC Central Committee and the State Council of the People's Republic of China attach great importance to the development and utilization of data factors. In 2015, the State Council issued the "Action Outline for Promoting Big Data Development," explicitly calling for accelerated government data openness and sharing. In 2021, the "Data Security Law of the People's Republic of China" and the "Personal Information Protection Law of the People's Republic of China" were promulgated successively, establishing the legal boundaries for data openness. In 2022, the "Opinions of the CPC Central Committee and the State Council on Building Data Basic Institutions to Better Leverage the Role of Data Factors" further designated public data as an important supply-side force in data factor market construction. In 2023, the National Data Bureau was officially established to coordinate the development, utilization, and open sharing of public data resources. Simultaneously, the national strategies of carbon peaking, carbon neutrality, and common prosperity have imposed dual requirements on urban development to balance green transformation and inclusive growth. Against this backdrop, whether public data openness can effectively drive urban inclusive green growth is of significant policy and practical importance.

3.2. Research Hypotheses

3.2.1. The Enabling Effect of Public Data Openness

Data factors possess the unique properties of non-rivalry and low marginal cost [42]; the same dataset can be reused simultaneously by multiple entities without diminishing in value. Public data open platforms, through the large-scale release of authoritative data in areas such as government services, economic management, and livelihood protection, exhibit significant positive externalities that can simultaneously act on both the inclusive growth and green growth dimensions.

In the inclusiveness dimension, public data open platforms help bridge the information divide between urban and rural areas and across different income groups by enhancing the transparency of public services and lowering the threshold for information access. Zhang and Xia [43] confirmed that public data openness significantly improves the level of basic public services by enhancing social inclusiveness, enabling middle-and low-income groups to more equally access public services such as education, healthcare, and social security more equally, thereby contributing to narrowing income gaps and equalizing basic public

services. Wang and Li [36] also found that public data openness exhibits inclusive and universally beneficial characteristics, capable of compensating for regional factor endowment gaps. In the green dimension, public data open platforms provide market entities with the data foundation needed for green decision-making by releasing high-value data such as environmental monitoring, corporate carbon emissions, and energy consumption statistics, driving enterprises to conserve energy and reduce emissions while strengthening the environmental oversight capacity of the public and regulatory agencies [32]. The high mobility of data enables real-time tracking of cross-regional pollution sources and precise accounting of carbon emissions, providing data support for regional joint prevention and control mechanisms.

Hypothesis 1. *Public data open platforms can promote urban inclusive green growth.*

3.2.2. The Mediating Effect of Digital Inclusive Finance

Digital inclusive finance is an important link connecting data openness with IGG. Traditional financial services in China face a significant "financial exclusion" dilemma: small and medium-sized enterprises (SMEs), rural residents, and low-income groups struggle to obtain equal financial services due to lack of credit records and information opacity [11], which not only exacerbates income distribution inequality but also constrains the accessibility of capital needed for green technology innovation. Digital inclusive finance, by reducing financial service costs and broadening the boundaries of financial services, helps bridge the aforementioned "financial divide" and has a direct promoting effect on inclusive green growth [44]. Public data open platforms provide the critical data foundation for the development of digital inclusive finance. Wang and Li [36] pointed out that public data openness, by supplying enterprise-related public data resources such as business registration, social security contributions, tax records, administrative penalties, and judicial proceedings, can effectively alleviate information asymmetry among banks, governments, and enterprises, improving financial service coverage and availability to build an inclusive and progressive financial ecosystem. Zhang [45] further confirmed that government data open platforms, by disclosing data on environmental penalties and carbon emissions, provide the information basis for financial institutions to identify green risks, promoting green finance development and channeling social capital toward green sectors.

Hypothesis 2. *Public data open platforms can promote the development of digital inclusive finance, thereby promoting inclusive green growth.*

3.2.3. The Mitigating Effect of Factor Market Distortion

Factor market distortion is an institutional barrier affecting IGG in Chinese cities. In the labor market, the household registration system and urban–rural segmentation impede cross-regional labor mobility, which not only constrains the optimal allocation of labor resources but also exacerbates urban–rural income disparities. In the capital market, financial resource misallocation and financing barriers cause SMEs and green industries to face financing constraints, inhibiting innovation-driven green growth. In the knowledge and technology factor market, information asymmetry leads to low efficiency in technology diffusion, widening the gap in green technology accumulation across cities [46]. These factor market distortions collectively form "institutional friction" that hinders inclusive green growth.

Public data open platforms mitigate factor market distortions through multiple mechanisms. First, public data openness reduces the degree of information asymmetry in market transactions [36,45], making factor price signals more accurate and helping guide factors

to agglomerate in high-efficiency green industries. Second, the opening of government data such as business registration, administrative penalties, and credit ratings can effectively improve the social credit system, reduce market transaction costs, and lower the institutional barriers to labor and capital factor mobility. Third, the government's real-time monitoring capability of market operations through data open platforms helps precisely identify and correct factor allocation inefficiencies, guiding resources toward green and efficient sectors. Zhao et al. [46] explicitly pointed out that the core mechanism through which the digital economy promotes inclusive growth lies in optimizing the allocation of labor, capital, and knowledge factors through cross-regional flows, building effective factor market logic consistent with the role of public data openness in alleviating factor distortion.

Hypothesis 3. *Public data open platforms can mitigate factor market distortions, thereby promoting inclusive green growth.*

3.2.4. The Enhancing Effect of Information Circulation

Information circulation efficiency is a fundamental condition affecting urban IGG. Sufficient information circulation helps reduce transaction costs, improve resource allocation, promote technology diffusion, and enhance market entities' awareness and ability to seize green development opportunities [32,43]. However, information asymmetry prevalent within and across cities creates "information friction" that impedes technology diffusion, market integration, and inclusive growth. This "friction" is particularly pronounced in less developed regions and is an important cause of the persistent widening of interregional development gaps [4].

Public data open platforms enhance information circulation efficiency through multiple channels. On the one hand, the opening of public data breaks down the "data silos" caused by administrative fragmentation, empowering cross-regional precision collaborative governance [32]. The same environmental or economic data can be used simultaneously by multiple entities, generating significant information spillover effects that effectively narrow the information access gap among various market entities. On the other hand, the establishment of public data platforms has expanded the breadth and depth of government information disclosure, enabling various market entities to more conveniently access policy dynamics, market opportunities, and green technology information, which helps stimulate innovation and entrepreneurial vitality and green investment willingness [33]. Cross-regional data circulation also promotes interregional technology cooperation and experience sharing, accelerating the diffusion of green innovation outcomes and providing technological channels for less developed regions to narrow the green development gap with advanced regions.

Hypothesis 4. *Public data open platforms can enhance urban information circulation levels, thereby promoting inclusive green growth.*

Based on the above hypothesis analysis, this paper constructs the theoretical framework of public data openness driving urban IGG, as illustrated in Figure 2. Public data openness operates through three parallel mechanisms: (1) reducing green financial information asymmetry and promoting digital inclusive finance development; (2) reducing institutional friction in factor markets and mitigating factor market distortions; and (3) breaking down data silos, diffusing green information, and enhancing urban information circulation levels. These mechanisms ultimately converge to form an overall enabling effect on urban IGG.

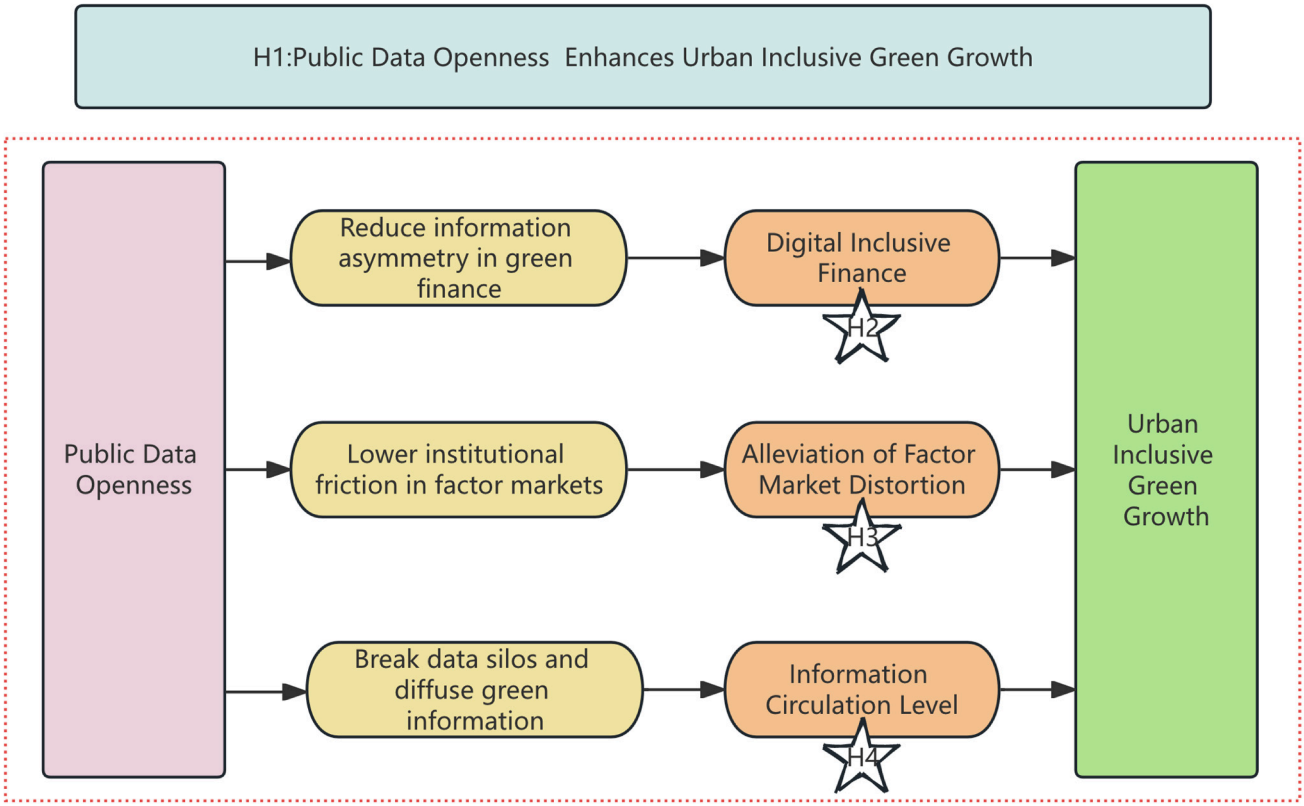


Figure 2. Transmission mechanism of public data open platforms.

4. Research Design

4.1. Model Specification

To overcome the model misspecification and overfitting problems inherent in traditional regression models, this paper draws on the research approaches of Zhang and Li [4] and Zhao et al. [46], employing the DDML framework to identify the causal effect of public data openness on urban IGG. Traditional linear regression models have limitations when handling high-dimensional control variables and complex nonlinear relationships, whereas DDML performs debiased estimation of the treatment and outcome variables separately through regularized machine learning methods, effectively mitigating regularization bias and overfitting problems.

Specifically, this paper constructs the following partially linear regression model:

$$Y_{it} = \theta_0 D_{it} + g_0(X_{it}) + U_{it} \quad (1)$$

$$E\{U_{it}|X_{it}\} = 0$$

where i denotes the city; t denotes the year; Y_{it} is the dependent variable, representing the inclusive green growth index of city i in year t; D_{it} is the treatment variable, a policy indicator for the "public data open platform" that takes the value of 1 if the platform has been launched and 0 otherwise; θ_0 is the core treatment effect parameter; X_{it} is the high-dimensional control variable set; U_{it} is the error term with a conditional mean of zero.

The primary reasons for adopting the DDML method are as follows. As a composite indicator, inclusive green growth is jointly influenced by numerous urban characteristic variables, and complex nonlinear interactions may exist between control variables and IGG. The linear functional form assumption of the traditional difference-in-differences method faces the "curse of dimensionality" and model misspecification risks. DDML, through orthogonalization, delegates the functional form of high-dimensional control variables to

machine learning algorithms for flexible fitting, effectively stripping away confounding factors without pre-specifying a concrete functional form, and obtaining $\sqrt{n}$-consistent estimates of the policy effect. Moreover, since the pilot selection of public data open platforms is not strictly random, the double robustness property of DDML renders it insensitive to model misspecification on both the propensity score and outcome regression sides, providing a more reliable estimation basis for causal inference.

4.1.1. DDML Implementation Details

In the partially linear model of Equation (1), the two nuisance functions, $m_0(X) = E(D|X)$ and $g_0(X) = E(Y|X)$, are estimated using random forests as the base learner. The learner is implemented through the pystacked command, which calls the random-forest regressor of the Python (version 3.9.6) scikit-learn library; because no manual tuning is applied, the default hyperparameters of the implementation are used: the number of regression trees is 100; the maximum depth of each tree is unrestricted; the minimum number of observations per leaf is 1; and all predictor variables are considered as candidate splitting variables at each split. Because random forests are a regularized ensemble method, no explicit L1/L2 penalty parameter is imposed; regularization is achieved through bagging and the ensemble-averaging structure itself.

Following Chernozhukov et al. [37], cross-fitting is implemented with K = 5 folds, corresponding to an 80:20 training-to-estimation split in each iteration: the nuisance functions are trained on four folds and used to predict the remaining fold, and the treatment effect parameter θ is then estimated by ordinary least squares regression of the residualized outcome on the residualized treatment variable. The reported estimates are aggregated across folds using the robust (median-based) statistics of the ddml package (version 1.4.4). All estimations are conducted in Stata 18 using the ddml package with pystacked, and the random seed is fixed at 42 to ensure full reproducibility.

4.1.2. Measurement of the Inclusive Green Growth Index

The IGG index is constructed with the entropy-weighting method. Let x_{ij} denote the raw value of secondary indicator j for city–year observation i (i = 1, . . ., n), with inverse indicators first converted by the reciprocal so that larger values always indicate better performance. All indicators are then rescaled to the unit interval by min–max normalization over the full sample and shifted by a constant c:

$$Z_{ij} = \frac{x'_{ij} - min_i x'_{ij}}{max_i x'_{ij} - min_i x'_{ij}} + c \quad (2)$$

where the extrema are taken over all city–year observations, and c = 0.001 is a uniform shift ensuring strictly positive values before the logarithm is taken; being strictly monotone, it preserves the ordering of all observations without truncation or winsorization.

The information entropy and the weight of indicator j are:

$$e_j = -\frac{1}{In\ n}\sum_{i=1}^{n} P_{ij} In P_{ij}\,, \quad W_j = \frac{(1 - e_j)}{\sum_j (1 - e_j)} \quad (3)$$

where $P_{ij} = \frac{Z_{ij}}{\sum_i Z_{ij}}$ is the share of observation i in indicator j. The composite index is obtained as:

$$IGG_i = \sum_{k=1}^{4} W_k S_{k,i} \quad (4)$$

where the same procedure is applied once within each dimension, and once across the four dimension scores, with W_k and $S_{k,i}$ the weight and score of dimension k. The index lies

in (0, 1) and enters the regressions in natural logarithms, so the negative logged minima follow mechanically from values bounded between 0 and 1 rather than from data errors.

4.2. Variable Selection

4.2.1. Dependent Variable: Urban Inclusive Green Growth (IGG)

Based on the evaluation system for core IGG measurement indicators [47] and referencing current research on comprehensive evaluation indicators for urban IGG [4,48], this paper constructs an indicator system from four dimensions-economic development(ED), social equity (SP), green production and consumption (GC), and environmental protection (EP)and employs the entropy method for comprehensive measurement. The economic development dimension includes per capita GDP, GDP growth rate, share of core digital industry assets, and employment rate in the secondary and tertiary sectors. The social equity dimension includes the urban–rural per capita income ratio, number of unemployment insurance enrollees, education expenditure, and pension insurance coverage rate. The green production and consumption dimension includes energy consumption per unit of GDP, energy conversion rate, industrial sulfur dioxide emissions, and industrial soot emissions. The environmental protection dimension includes annual average PM2.5 concentration, green coverage rate of built-up areas, sewage and waste treatment rate, and industrial sulfur dioxide removal rate. After positive processing of negative indicators, the entropy method is used to obtain the composite index and sub-dimensional scores, as shown in Table 1.

Table 1. Comprehensive evaluation system of urban inclusive green growth.

Primary Indicator	Secondary Indicator	Nature	Unit
Economic Development (ED)	Per capita GDP	+	CNY/person
	GDP growth rate	+	%
	Share of core assets of digital industries	+	%
	Employment share of secondary and tertiary industries	+	%
Social Equity (SP)	Urban–rural per capita income ratio	−	%
	Number of unemployment insurance participants	+	10,000 persons
	Education expenditure	+	100 million CNY
	Pension insurance coverage rate	+	%
Green Production and Consumption (GC)	Energy consumption per unit of GDP	−	tce/10,000 CNY
	Energy conversion rate	+	%
	Industrial SO_2 emissions	−	10,000 tons
	Industrial soot emissions	−	10,000 tons
Environmental Protection (EP)	Annual average PM2.5 concentration	−	$\mu g/m^3$
	Green coverage rate of built-up areas	+	%
	Sewage and waste treatment rate	+	%
	Industrial SO_2 removal rate	+	%

Notes: "+" denotes a positive indicator; "−" denotes a negative indicator.

4.2.2. Treatment Variable: Public Data Open Platform Policy Dummy Variable

In 2012, China's first local public data open platform was launched, after which cities successively followed suit with significant differences in launch timing. This progressive rollout provides an ideal quasi-natural experimental setting: policy implementation has strong exogeneity, and the staggered launch times across cities create natural treatment and control groups. Based on this, the treatment variable is constructed according to whether a city has established a public data open platform, taking the value of 1 for the year of and years following platform launch, and 0 for years prior to launch.

4.2.3. Control Variables and Mechanism Variables

Considering that urban inclusive green growth is jointly driven by multiple factors, the following control variables are selected: (1) Education development level, measured as the ratio of local government education expenditure to GDP, as education investment is the foundation of human capital accumulation and technological innovation. (2) Fixed asset investment, measured as the share of total social fixed asset investment in GDP, as capital accumulation significantly affects industrial structure and energy consumption patterns. (3) Government intervention, measured as the share of fiscal expenditure in GDP, as local government fiscal behavior plays an important guiding role in resource allocation and environmental protection. (4) Government science and technology expenditure, measured as the share of S&T expenditure in GDP, as technological innovation is a key engine for promoting green development. (5) Internet penetration rate, measured as the share of broadband subscribers in the total population, as digital infrastructure is an important prerequisite for public data openness to take effect. (6) Openness to the outside world, measured as the share of total import and export value in GDP, as trade openness influences urban green growth performance through technology spillover and competition effects. (7) Population density, measured as the logarithm of the ratio of year-end population to administrative area, as population agglomeration can generate economies of scale but may also intensify resource and environmental pressures. (8) Resource and environment, measured as the share of environmental pollution control investment in GDP, as environmental governance investment directly relates to the material foundation for urban green development. (9) Human capital, measured as the logarithm of the number of university students per 10,000 persons, as high-quality talent is the fundamental guarantee for achieving inclusive growth. For mechanism variables, this paper selects the digital inclusive finance composite index, factor market distortion index, and total postal business volume to reveal the pathways through which public data openness affects urban IGG from the three dimensions of financial inclusion, resource allocation optimization, and information circulation enhancement.

4.2.4. Data Description and Descriptive Statistics

This paper selects prefecture-level and above cities in China from 2010 to 2024 as the research sample. Data are primarily sourced from the China City Statistical Yearbook, China Regional Economic Statistical Yearbook, China Environmental Statistical Yearbook, China Science and Technology Statistical Yearbook, China Urban Construction Statistical Yearbook, provincial and municipal statistical yearbooks and bulletins, and government open data. The launch dates of public data open platforms were manually compiled from city government data open platforms, government portals, and relevant policy documents. For a small number of missing values, linear interpolation was used for supplementation; price-related indicators were deflated using 2010 as the base period.

To mitigate the influence of variable distribution skewness, heteroscedasticity, and scale differences on estimation results, the natural logarithm is taken of the dependent variable, control variables, and continuous mechanism variables in the empirical analysis. The core explanatory variable (Data) is a policy dummy variable and is not log-transformed. Therefore, the occurrence of negative values for some variables in the descriptive statistics table does not indicate data anomalies but rather results from the natural logarithm of ratio-type or index-type variables with values less than 1.

Table 2 reports the descriptive statistics of the main variables. The mean of the public data open platform variable is 0.236, indicating that approximately 23.6% of the city–year observations in the sample correspond to the period after platform launch. The main control and mechanism variables exhibit a certain degree of dispersion, indicating

significant differences across cities in terms of green development foundation, digital infrastructure, fiscal S&T investment, human capital accumulation, marketization level, and information circulation capacity, providing a data foundation for subsequent baseline regression, mechanism testing, and heterogeneity analysis.

Table 2. Descriptive statistics of main variables.

Variable	Definition	Mean	SD	Min	Max
IGG	Log of IGG index	−1.128	0.327	−2.572	−0.082
Public Data Open Platform	Whether the city has launched a platform	0.236	0.424	0.000	1.000
Education Expenditure	Log of education expenditure level	−3.443	0.497	−6.149	0.217
Fixed Asset Investment	Log of fixed asset investment level	3.224	0.763	−4.670	6.311
S&T Expenditure	Log of S&T expenditure level	−5.957	1.050	−10.095	0.424
Fiscal Expenditure	Log of fiscal expenditure level	4.499	0.976	1.576	8.061
Internet Infrastructure	Log of broadband access level	−1.534	0.841	−5.853	6.109
Openness	Log of openness level	−7.324	1.753	−17.420	−3.128
Population Density	Log of population density	−3.605	1.211	−9.005	0.813
Environmental Governance	Log of environmental governance investment	0.279	0.228	−0.028	1.519
Human Capital	Log of university students per 10,000 persons	1.463	0.356	−0.367	2.062
Digital Inclusive Finance	Log of digital inclusive finance level	5.241	0.528	2.891	6.002
Marketization	Log of factor market distortion index	0.447	0.052	0.248	0.542
Information Circulation	Log of total postal business volume	11.068	1.530	5.303	17.098

5. Empirical Analysis

5.1. Baseline Regression

This paper employs the DDML method to estimate the causal effect of public data open platform construction on urban inclusive green growth. The advantage of the DDML method lies in its use of machine learning algorithms for non-parametric fitting and residualization of high-dimensional control variables, effectively overcoming regularization bias while relaxing traditional linear parametric assumptions, to achieve consistent estimation of the treatment effect. The sample split ratio is 1:4, and the regression results are presented in Table 3. Models (1) and (2) report the baseline regression results, where Model (1) controls for city fixed effects, time fixed effects, and the linear terms of control variables within the sample period. It can be observed that the coefficient of the public data open platform policy on urban IGG is significant at the 1% level, indicating that the launch of public data open platforms can enhance the level of urban inclusive green growth. Model (2) further controls for the quadratic terms of control variables on the basis of Model (1), and the regression result remains significant at the 1% level.

Table 3. Baseline regression results of double machine learning.

Variable	IGG (1)	IGG (2)	ED (3)	SP (4)	GC (5)	EP (6)
Data	0.0223 *** (0.005)	0.0222 *** (0.005)	0.0138 *** (0.005)	0.0470 *** (0.011)	0.0245 *** (0.008)	0.0165 *** (0.004)
N	4395	4395	4395	4395	4395	4395
Linear controls	Yes	Yes	Yes	Yes	Yes	Yes
Quadratic controls	No	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes

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

Models (3)–(6) further examine the dimension-specific effects of public data open platforms on urban IGG. The estimated coefficients for economic development, social equity, green production and consumption, and environmental protection are all positive and statistically significant at the 1% level. These findings indicate that public data openness is associated with broad-based improvements across the economic, social, production, and environmental aspects of inclusive green growth.

With respect to social equity, public data openness can reduce information-search costs and lower access barriers to public services. The disclosure of information related to education, healthcare, social security, employment, and other public services may enable a broader range of residents and enterprises to access policy information and public resources more efficiently, thereby supporting a more inclusive allocation of development opportunities.

For green production and consumption, the availability of environmental monitoring information, green standards, energy-use data, and corporate environmental records can improve information transparency. Such information may encourage firms to adopt cleaner production practices and assist consumers in making more environmentally conscious choices. Through these channels, public data openness can support the diffusion of green technologies and the development of green consumption patterns.

Regarding environmental protection, public data openness can enhance the transparency of environmental regulation and strengthen public oversight. Timely access to pollution, emissions, and environmental-quality information may help government agencies, firms, and the public identify environmental risks and improve the targeting of environmental governance measures. However, improvements in environmental quality may also depend on complementary investments, technological adoption, and the implementation capacity of local governments.

Finally, public data openness can contribute to economic development by improving the urban information environment and reducing transaction frictions. Greater access to market, credit, regulatory, and industrial information may facilitate resource matching, entrepreneurial decision-making, and innovation activities. The dimension-specific estimates should therefore be interpreted as complementary evidence that public data openness generates multidimensional policy benefits, rather than as establishing a statistical ranking of effect magnitude across the four dimensions.

5.2. Robustness Checks

5.2.1. Parallel-Trend Validation

To assess the conditional parallel-trends assumption, we conduct a dynamic event-study analysis within the DDML residualization framework. The outcome and event-time treatment indicators are residualized with respect to covariates, city fixed effects, and year fixed effects using cross-fitted random forests. The year before platform launch is used as the reference period.

As shown in Figure 3, the coefficients for the third and second years before platform launch are statistically indistinguishable from zero. The joint test yields $\chi^2(2) = 2.40$ ($p = 0.3007$), failing to reject the null hypothesis that the two pre-treatment coefficients are jointly zero. Therefore, within the common-support sample, we find no evidence of systematic differential pre-treatment trends between treated and control cities. The post-treatment coefficients are positive, consistent with the baseline DDML estimates.

5.2.2. Narrowing the Sample Time Window

The construction of public data open platforms by Chinese local governments exhibits significant phased characteristics. The period from 2012 to 2015 was an exploratory start-up phase, during which only a few economically developed cities pioneered platform estab-

lishment, and the policy implementation intensity, data openness scope, and subsequent batches of early-treated cities differed qualitatively. After 2016, public data openness entered a phase of comprehensive rollout, with a large number of cities establishing platforms during this period, the share of treated cities increasing significantly, and the identification conditions for treatment effects improving substantially. The sample is therefore restricted to 2016–2024 for regression. The results, shown in Table 4 (1), indicate that the regression result remains significant at the 1% level.

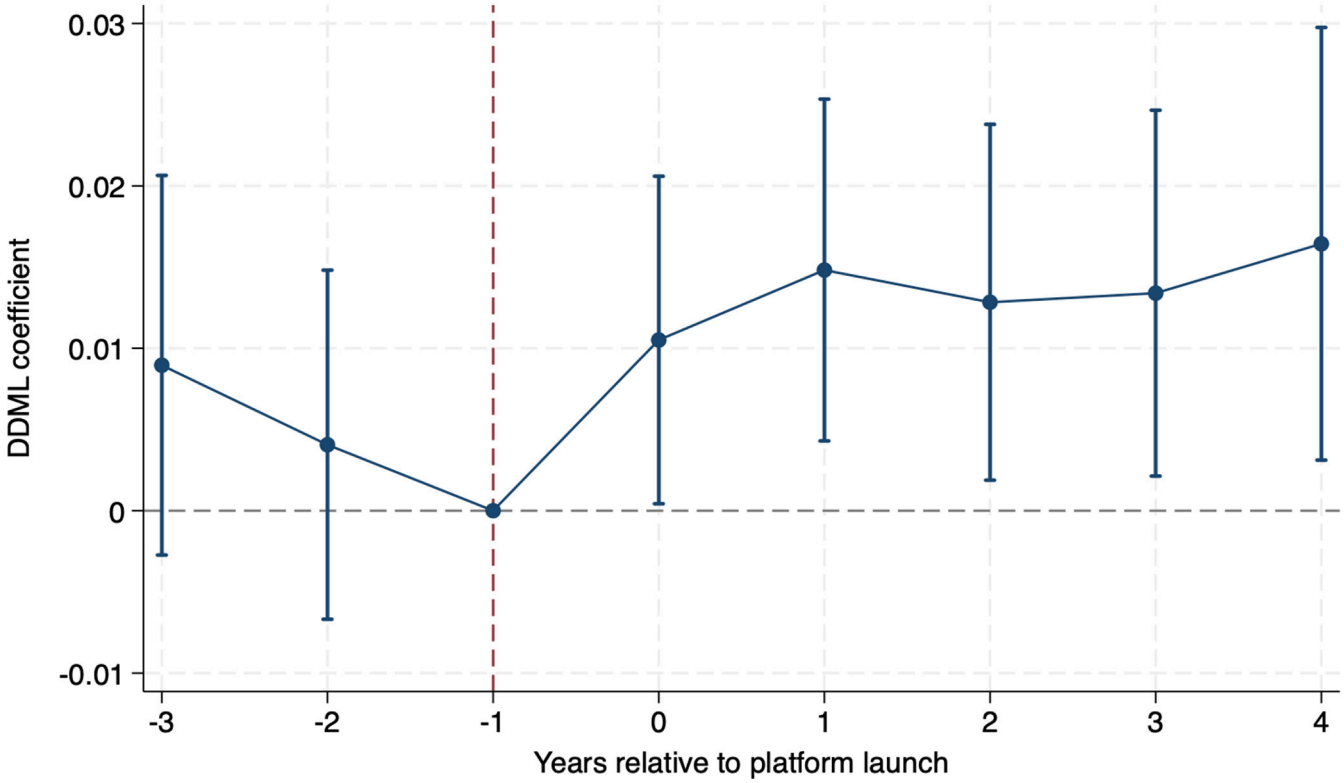


Figure 3. Dynamic DDML estimates of the effect of public data open platforms on inclusive green growth. Notes: The figure reports event-time coefficients estimated using the DDML residualization framework. The year before platform launch (k = −1) is the reference period. Error bars denote 95% confidence intervals. The event-study sample retains treated cohorts with at least three pre-treatment and four post-treatment observations; never-treated cities are retained as controls.

Table 4. Robustness Test Results.

Variable	(1) Narrowed	(2) Excl. Early	(3) K = 3	(4) K = 8	(5) GBM
Data	0.0155 *** (0.004)	0.0168 *** (0.005)	0.0240 *** (0.005)	0.0212 *** (0.005)	0.0500 *** (0.004)
N	2625	4215	4395	4395	4395
First-order controls	Yes	Yes	Yes	Yes	Yes
Second-order controls	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes

Note: Heteroscedasticity-robust standard errors in parentheses; ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

5.2.3. Removing Early Pilot Cities

Early pilot cities are significantly ahead of the vast majority of cities in terms of platform construction timing. These cities typically have higher fiscal revenues and per

capita GDP levels, stronger information technology infrastructure, more sophisticated governmental transparency institutional arrangements, and policy pilot authority granted by higher-level governments. Therefore, removing these sample cities for regression yields results as shown in Table 4 (2), which remain significant at the 1% level, demonstrating the robustness of the baseline regression results.

5.2.4. Changing the Cross-Fitting Fold Number

Cross-fitting is the core technical innovation of the DDML framework. It partitions the sample into K disjoint subsets, alternately using $K-1$ folds for training and the remaining one fold for residualized estimation, thereby effectively eliminating the "own-regression" bias introduced by the overfitting of machine learning algorithms [37]. The choice of the number of cross-fitting folds K directly determines the sample ratio between the training set and the estimation set, thereby affecting the statistical properties of the estimator. Therefore, the sample split ratio is changed from the original 1:4 to 1:2 and 1:7 to explore the impact on the research conclusions. The results are shown in Table 4 (3) and (4). When k varies within the range of 3–8, the core estimated coefficient remains stable, indicating strong robustness to the cross-fitting fold setting.

5.2.5. Changing the Machine Learning Method

A key advantage of the DDML framework is its "Neyman orthogonality" to the choice of base learners—in theory, as long as the base learner converges to the conditional expectation function at a sufficiently fast rate, the final treatment effect estimate should be insensitive to the specific algorithm choice. However, in finite samples, differences in the inductive bias and regularization mechanisms of different algorithms may lead to substantive divergence in estimation results. Therefore, the base learner algorithm is replaced with gradient boosting for robustness checks. The study by Yang et al. [38] provides a methodological basis for the application of the gradient boosting algorithm within the DDML framework. The results are shown in Table 4 (5). The coefficient increases substantially compared to the RF baseline, reflecting that GBM, through more thorough residualization, releases a more complete treatment effect. However, the sign direction, significance level, and economic implications remain unchanged, and the core conclusion holds under algorithm substitution.

The collective results of the five robustness checks demonstrate that the Data coefficient is positive in all tests and significant at the 1% level. Although the coefficient magnitudes exhibit moderate fluctuations, no sign reversal or loss of significance occurs. Changes in sample screening, cross-fitting parameters, and machine learning algorithms do not overturn the baseline conclusion. Therefore, the baseline regression results have a solid empirical foundation.

6. Transmission Mechanisms and Heterogeneity Analysis

6.1. Transmission Mechanism Tests

Referring to the discussion of mechanism testing in causal inference by Jiang [49], this paper adopts a two-step DDML framework to identify potential transmission mechanisms. The first step tests whether public data open platforms significantly affect the mechanism variable; the second step adds the mechanism variable to the baseline model and observes the magnitude of the decrease in the Data coefficient relative to the baseline model. The mediation proportion is estimated as "(baseline Data coefficient minus Data coefficient after adding the mechanism variable) divided by the baseline Data coefficient." Two caveats should be noted. First, this method is primarily used to reveal empirical facts related to mechanisms and does not equate to a complete decomposition of the causal effect. Second,

because each mechanism variable is introduced individually in these two-step regressions, the resulting proportions capture the relative importance of each channel but overlap across channels and therefore cannot simply be summed.

The results are shown in Table 5. When the three channels are evaluated individually, digital inclusive finance and factor market distortion each account for approximately 50% of the total policy effect, while the information circulation channel accounts for approximately 18%. To assess the joint explanatory power of the three channels and the degree of overlap among them, we additionally estimate a joint specification that includes all three mechanism variables simultaneously (Table 6); to ensure comparability, both the baseline and the joint specification are estimated on the common sample of 4090 city–year observations with complete mediator data. In this specification, they are both significant at the 1% level, implying that the three channels jointly transmit approximately 30% of the total effect, with the remainder operating through channels not yet captured by our analysis. Notably, the jointly transmitted share (≈30%) is smaller not only than the sum of the individually estimated shares (≈118%) but also than the two largest individual shares (≈50%) themselves. This pattern reflects the overlap across the three channels: because the mediators are correlated, each single-mediator specification partially absorbs variation associated with the other channels, so the individually estimated shares embed double-counting, whereas the joint specification removes this overlap and therefore provides a conservative measure of the joint transmission. These results confirm that substantial overlap exists across the three channels and that they operate under a common underlying logic rather than as independent, additive pathways. The three pathways should therefore be understood as complementary rather than mutually exclusive transmission mechanisms. Digital inclusive finance and factor market distortion both point to the core mechanism of reducing information asymmetry. The effect of the information circulation channel is relatively limited, reflecting that the governance dividends of public data openness are realized more through digital information circulation than physical information transmission. However, the statistical significance of this channel indicates that improvements in physical information infrastructure remain an indispensable foundational condition for the implementation of digital governance effects. The essence of public data openness enhancing government governance lies in using the institutional openness of the information supply side to break down information barriers. Whether it is the scarcity of financial credit data, the distortion of factor price signals, or the restriction of information circulation channels, the root cause is market failure and governance inefficiency resulting from information asymmetry. Data openness precisely provides multiple and complementary transmission pathways for governance improvement from the perspective of information democratization.

Table 5. Mechanism test results.

Variable	Digital Inclusive Finance		Factor Market Distortion		Postal Business Volume	
	Step 1	**Step 2**	**Step 1**	**Step 2**	**Step 1**	**Step 2**
Data	0.039 *** (0.007)	0.011 ** (0.004)	−0.002 *** (0.001)	0.011 ** (0.005)	0.09 *** (0.031)	0.018 *** (0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	4090	4090	4395	4395	4395	4395

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

Table 6. Joint mechanism test results.

	(1)	(2)
	Baseline	**Joint Specification**
Data	0.0206 ***	0.0144 ***
	(0.0045)	(0.0045)
Controls	Yes	Yes
City FE	Yes	Yes
Year FE	Yes	Yes
N	4090	4090

Note: Heteroskedasticity-robust standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. DDML estimates use random-forest nuisance functions implemented through pystacked, 5-fold cross-fitting, and seed 42. Column (2) includes digital inclusive finance, factor market distortion, and postal business volume simultaneously.

6.1.1. Promoting Digital Inclusive Finance Development

Public data openness significantly promotes the development of digital inclusive finance. The first-step estimate shows that the effect of public data openness on the digital inclusive finance index is 0.039, significant at the 1% level, indicating that the data openness policy provides financial institutions with richer public information resources, broadens the data foundation for credit assessment, and lowers the threshold and cost of financial service penetration, thereby driving improvements in both the breadth of coverage and depth of usage of digital inclusive finance. The second-step estimate shows that after controlling for the digital inclusive finance index, the effect of public data openness on the IGG index decreases from the baseline 0.022 to 0.011, a coefficient reduction of approximately 50%. This result indicates that digital inclusive finance is an important transmission channel through which public data openness enhances government governance: data openness $\rightarrow$ enriching financial credit data $\rightarrow$ promoting inclusive finance penetration $\rightarrow$ enhancing financial inclusiveness and service availability $\rightarrow$ improving government governance performance. The development of digital inclusive finance not only directly improves the financial accessibility of residents and enterprises but also indirectly promotes local economic vitality and governance effectiveness by reducing information asymmetry and alleviating financing constraints. Approximately half of the policy effect is realized through this channel.

6.1.2. Mitigating Factor Market Distortion

Public data openness significantly mitigates factor market distortion. The factor market distortion index, following the method of Sun and Lin [50], is calculated based on the Cobb–Douglas production function as the ratio of the marginal product of capital and labor factors to their market prices, to measure the degree of deviation of factor prices from marginal products. A larger index indicates more severe deviation between market prices and marginal products of factors, i.e., a higher degree of factor misallocation. The first-step estimate shows that the effect of public data openness on the factor market distortion index is -0.002, significant at the 1% level, with a negative coefficient, indicating that the data openness policy effectively reduces the degree of price distortion in factor markets. The opening of public data enables market entities to access factor price information more timely and comprehensively, reducing factor misallocation caused by information barriers and promoting more efficient flows of labor, capital, and other factors across sectors and regions. The second-step estimate shows that after controlling for factor market distortion, the effect of public data openness on the IGG index decreases to 0.011, a coefficient reduc-

tion of approximately 50%. This result reveals the governance transmission mechanism of optimal factor allocation: data openness → reducing factor price information asymmetry → mitigating factor market distortion → improving resource allocation efficiency → enhancing government governance performance. The mitigation of factor market distortion signifies more rational resource allocation, improved output efficiency of the local economy, and an improved institutional environment for government governance. This channel similarly accounts for approximately half of the total policy effect.

6.1.3. Enhancing Urban Information Circulation

Public data open platforms have a significant promoting effect on regional postal business volume, with a coefficient of 0.09, indicating that data openness reduces information asymmetry and activates the interaction between logistics and information flows. After further incorporating postal business volume into the baseline model, the estimated coefficient of Data on IGG is 0.018, still significantly positive. This result confirms that the vitality of the postal and logistics system is an effective transmission pathway through which public data openness promotes IGG. Data openness facilitates rural e-commerce, green product circulation, and regional market integration, thereby promoting growth while balancing inclusiveness and environmental sustainability.

6.2. Heterogeneity Analysis

The preceding baseline regression and mechanism tests have revealed the average treatment effect of public data open platforms on urban IGG and its transmission channels. However, Chinese cities exhibit significant endowment differences in economic scale, human capital accumulation, and institutional foundation, and policy effects may be nonuniformly distributed across cities. This section conducts heterogeneity analysis from three dimensions: urban economic scale, human capital level, and initial IGG level. The results are shown in Table 7.

Table 7. Heterogeneity test results.

Variable	(1) City Scale			(2) Human Capital			(3) Initial IGG		
	Small	**Medium**	**Large**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Data	0.023 ** (0.010)	0.021 *** (0.007)	0.030 *** (0.007)	0.026 *** (0.008)	0.039 *** (0.010)	0.065 *** (0.009)	0.021 ** (0.010)	0.025 *** (0.008)	0.048 *** (0.008)
N	1464	1461	1470	1479	1463	1453	1461	1469	1465
Linear controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quadratic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

6.2.1. Urban Economic Scale Heterogeneity

Sample cities are divided into small, medium, and large city groups based on per 630 capita GDP. The grouped DDML estimation results are shown in Column (1) of Table 7. The small city group has $\beta = 0.023$ (SE = 0.010, $p < 0.05$), the medium city group has $\beta = 0.021$ (SE = 0.007, $p < 0.01$), and the large city group has $\beta = 0.030$ (SE = 0.007, $p < 0.01$). The estimated coefficients for all three city types are significantly positive, but the effect magnitudes exhibit a clear gradient: the large city effect (0.030) is approximately 30% higher than that of small cities (0.023), with the medium city effect (0.021) in between. The spatial distribution is shown in Figure 3. Cities with larger economic scales typically possess more sophisticated information infrastructure, more active market transaction

ecosystems, and richer data application scenarios, giving them stronger capacity to absorb and convert the "supply-side shock" of public data openness. Specifically, large cities have a greater number of enterprise entities and broader industry coverage, enabling them to more efficiently embed open data into production decisions, supply chain management, and green innovation processes. Simultaneously, the digital governance capability and data operation and maintenance experience of large city governments are more mature, with the data update frequency, interface availability, and data quality of open platforms all superior to those of small and medium cities, allowing for a more sufficient release of policy dividends. Although small cities also benefit from data openness, the marginal release of policy effects is somewhat constrained by smaller market scale, insufficient data application demand, and weak digital infrastructure.

6.2.2. Human Capital Level Heterogeneity

The sample is divided into low, medium, and high human capital groups based on the logarithm of the number of university students per 10,000 persons. The results are shown in Column (2) of Table 7. The estimated coefficients for the three groups exhibit a strictly monotonic increasing trend: the low human capital group has $\beta = 0.026$ (SE = 0.008, $p < 0.01$), the medium human capital group has $\beta = 0.039$ (SE = 0.010, $p < 0.01$), and the high human capital group has $\beta = 0.065$ (SE = 0.009, $p < 0.01$). The spatial distribution is shown in Figure 4. From low to high human capital, the policy effect increases by approximately 2.5 times ($0.026 \rightarrow 0.065$), an increase far exceeding the gradient difference in the city scale dimension. On the one hand, public data openness releases unprocessed raw data resources, and the degree to which their governance and economic value is realized depends on the ability of market entities to interpret, clean, and analyze data. High human capital cities possess more abundant statistical analysis talent, data science practitioners, and technology transfer intermediaries, enabling them to more efficiently convert open data into actionable business insights and policy recommendations, thereby fully releasing the potential information value of data. Low human capital cities face a structural contradiction of "sufficient data supply but insufficient interpretation capacity," where the information content of open data is not fully mined, limiting policy effects to a lower level. On the other hand, human capital is a fundamental input factor for green innovation and digital technology application. Workers in high human capital cities more readily master energy-saving and emission-reduction technologies, digital production tools, and the information processing capabilities needed for green consumption decisions, enabling the governance dividends transmitted through the digital inclusive finance and information circulation channels to obtain more sufficient reception on the demand side.

6.2.3. Initial IGG Level Heterogeneity

The sample is divided into low, medium, and high initial IGG groups based on the 2012 IGG index. The results are shown in Column (3) of Table 7. The estimated coefficients for the three groups similarly exhibit a monotonic increasing pattern: the low initial IGG group has $\beta = 0.021$ (SE = 0.010, $p < 0.05$), the medium initial IGG group has $\beta = 0.025$ (SE = 0.008, $p < 0.01$), and the high initial IGG group has $\beta = 0.048$ (SE = 0.008, $p < 0.01$). The policy effect for the high initial IGG group (0.048) is approximately 2.3 times that of the low initial IGG group (0.021). The spatial distribution is shown in Figure 4. Cities with higher IGG levels typically indicate that they have established relatively mature institutional arrangements and market order across the four dimensions of economic development efficiency, social equity protection, green production modes, and environmental governance capacity. Public data openness, as an institutional information supply, requires embedding within existing institutional frameworks and governance ecosystems for its governance effectiveness to

be realized. The compliant use of open data requires sound privacy protection and data security institutions, data-driven public service optimization requires well-developed government informatization systems, and he environmental regulatory application of green data requires effective enforcement and accountability mechanisms. Cities with higher initial IGG levels possess more complete complementary institutional conditions, enabling public data openness to form positive synergies with existing governance systems, amplifying policy effects across multiple dimensions. Cities with lower initial IGG levels have weaker governance foundations, and data openness faces an institutional vacuum. Even when data is made available to society, the lack of effective institutional channels to convert it into tangible governance improvements constrains the release of policy dividends.

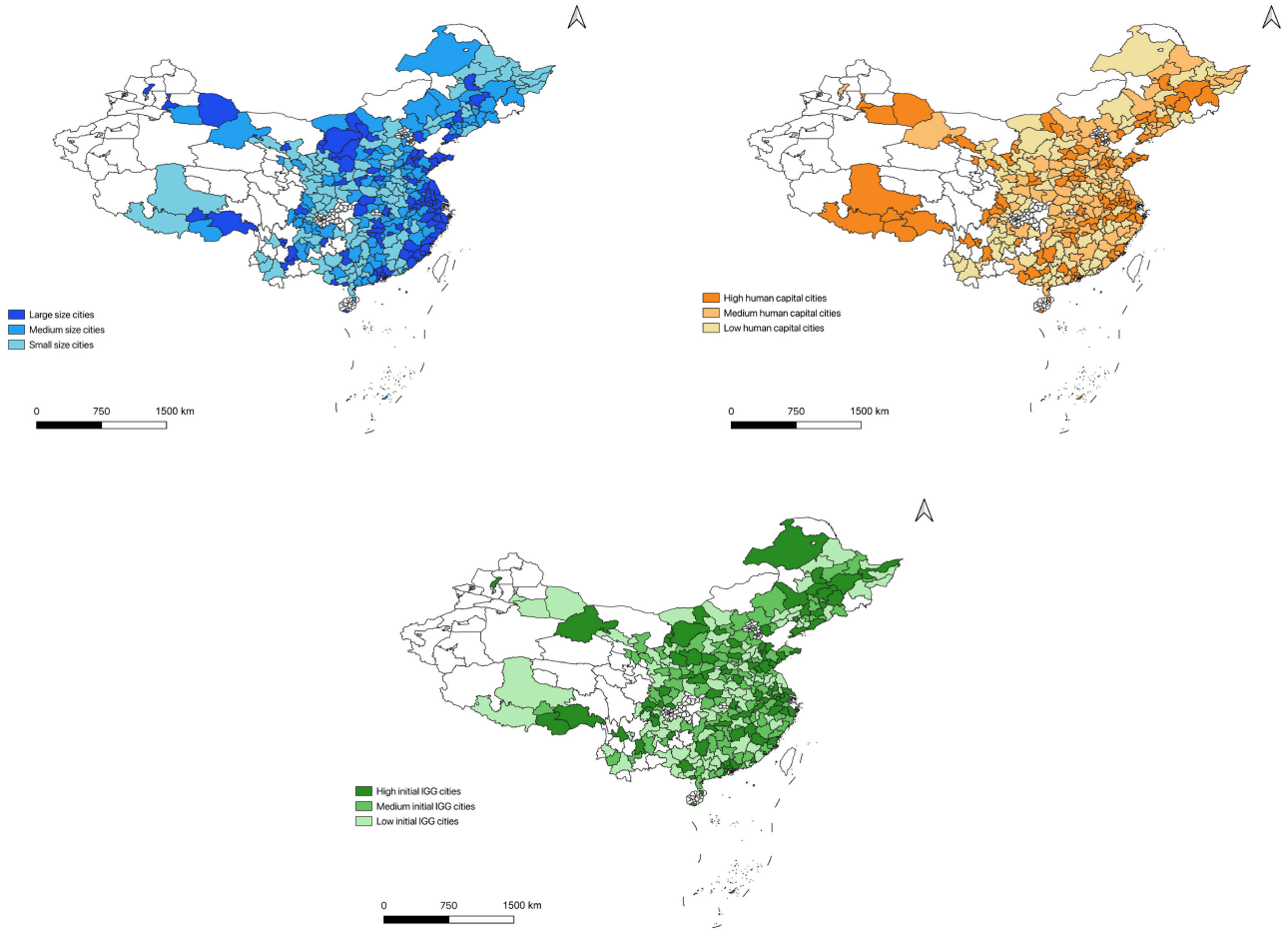


Figure 4. Heterogeneity analysis.

6.2.4. Summary

From a theoretical perspective, the heterogeneity patterns across the three dimensions collectively point to a core proposition: the value realization of data factors is subject to significant "complementary condition constraints." Although data itself possesses the technical characteristics of non-rivalry and low marginal cost, the release of its governance effects depends on three types of complementary conditions: the data interpretation capacity determined by the human capital of market entities, the data application demand supported by urban economic scale, and the reception capacity at the initial IGG level through institutional channels. This finding extends the explanatory framework of traditional factor endowment theory: in the era of data factors, urban development gaps stem not only from differences in the stock of physical capital and labor but also from gaps in the complementary factors needed to convert data factors into governance improvements.

Merely increasing data supply without cultivating complementary factors may lead to a predicament of "data abundance but governance poverty."

From a policy perspective, the heterogeneity analysis provides clear empirical evidence for advancing public data open platform construction in a locally adapted manner. For high-endowment cities, the focus should be on deepening the breadth and quality of data openness to release the deep governance value of data. For low-endowment cities, data openness should be advanced in coordination with human capital cultivation, infrastructure improvement, and institutional capacity building to avoid imbalanced distribution of policy dividends. Specifically, low-human-capital cities should simultaneously strengthen data science talent cultivation and digital literacy promotion; low-economic-scale cities should prioritize building the infrastructure support for data applications; and low-initial-IGG cities should focus on improving the institutional linkage between data openness and governance improvement, so that the governance dividends of data factors can be reasonably released under different endowment conditions.

7. Regional Heterogeneity Analysis

The heterogeneity analysis in Section 6.2 is confined to city-level dimensions. To examine whether the effect of public data openness also varies systematically with geographic location, this section re-estimates the baseline DDML specification on subsamples of eastern, central, and western cities. Figure 5 displays the geographic distribution of sample cities across the three regions: the eastern region mainly encompasses developed coastal provinces, the central region extends in a belt from north to south, and the western region occupies a vast inland hinterland. The grouped estimation results are reported in Table 8. The impact effects are not random; they are rooted in China's gradient regional strategy of "eastern region taking the lead—western development—central region rise," with different regions exhibiting systematic differences in resource endowments, industrial foundations, and data infrastructure construction. Table 8 illustrates the impact effects on the eastern, central, and western regions.

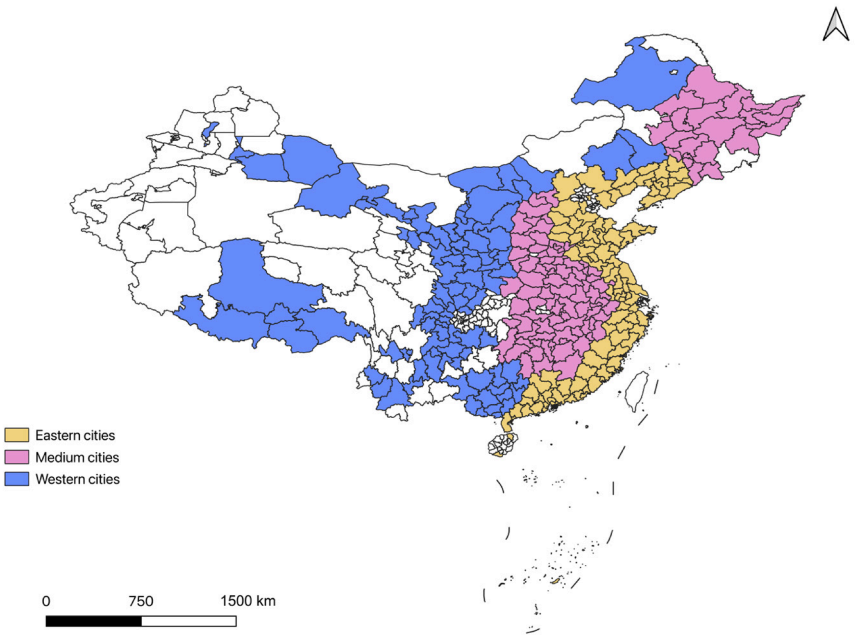


Figure 5. Regional heterogeneity of public data openness effects: Eastern, Central, and Western regions.

Table 8. Eastern–Central–Western heterogeneity test.

Variable	Eastern	Central	Western
Data	0.051 *** (0.009)	0.015 ** (0.007)	0.036 *** (0.010)
N	1479	1486	1400
Linear controls	Yes	Yes	Yes
Quadratic controls	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes

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

The policy effect in the eastern region is far higher than in the central and western regions, which is consistent with the logic of the heterogeneity patterns in the aforementioned city scale and human capital dimensions. The comprehensive endowment advantages of eastern cities in terms of economic scale, industrial structure, human capital accumulation, and digital infrastructure give them stronger conversion capacity for data openness policies. More critically, eastern cities possess more active market ecosystems and denser enterprise entities, enabling the information supply released by public data openness to be efficiently embedded in diverse governance scenarios such as financial credit assessment, factor price matching, supply chain optimization, and green innovation, thereby significantly amplifying the marginal contribution of policy effects.

The central region exhibits the smallest effect. This "central collapse" phenomenon can be understood at two levels. First, central cities occupy a transitional position between the eastern and western regions, lacking both the economies of scale and data application density of eastern cities and the unique policy preference effects enjoyed by western cities. In the national regional development strategy, western cities have long benefited from fiscal transfer payments, dedicated infrastructure investment, and talent introduction policies, among other multidimensional resource preferences. These complementary policies form institutional complementarities with public data openness, amplifying the governance effectiveness of data factors. Central cities, however, are relatively "marginalized" in policy resource allocation, and data openness lacks sufficient complementary condition support, constraining the release of policy effects. Second, the industrial structure of central cities is dominated by traditional manufacturing and resource-based industries, with a lower demand elasticity for data factors than the service- and innovation-economy-dominated industrial structure of the eastern region, resulting in relatively limited commercial conversion and governance application scenarios for open data.

The western region's effect exceeds that of the central region. This "western surpassing central" non-monotonic pattern reveals the important role of policy preferences and catch-up effects. Although western cities are inferior to central cities in terms of economic scale and human capital, as key beneficiaries of the national coordinated regional development strategy, they have received more central policy support, technical assistance, and financial subsidies during the construction of public data open platforms. The injection of these external institutional resources partially compensates for the deficiencies in western cities' own endowment conditions, enabling the implementation intensity and coverage quality of the data openness policy in western cities to surpass those of some central cities, generating positive superposition results. Moreover, due to the lower starting point of governance foundations in western cities, the marginal improvement space brought by data openness is larger. The marginal utility of information supply transitioning from extreme scarcity to basic availability far exceeds the incremental improvement from basic adequacy to

greater richness, forming a governance return increasing characteristic similar to a "catch-up effect."

8. Discussion

This paper is the first to incorporate public data open platforms and IGG into a unified causal inference framework. It is worth noting that existing literature has constructed IGG indices from multiple dimensions: Gu and Sun [8] measured Chinese urban IGG from three dimensions of economic development, social livelihood, and green sustainability; Han and Li [48] approached from three dimensions of economic growth, social equity, and ecological friendliness; and Zhang and Li [4] constructed an indicator system from four dimensions of economic growth, income distribution, welfare inclusiveness, and environmental pollution reduction. Building on this foundation, the present paper further separates the environmental dimension into two independent dimensions: "green production and consumption" and "environmental protection." The former characterizes the degree of green transformation in production processes and consumption behaviors, while the latter reflects the terminal state of environmental quality. This dimensional refinement enables the paper to distinguish the differential effects of public data openness at the "process end" and "outcome end" of green transformation, remedying the limitation of existing literature that mixes the two and cannot identify policy leverage points. The four-dimensional decomposition in this paper provides a finer analytical perspective for this field.

At the mechanism level, when all three mediators are entered jointly, the estimated policy effect declines from 0.0206 to 0.0144 and remains significant at the 1% level, indicating that the three channels jointly transmit approximately 30% of the total policy effect (Table 6). Notably, this joint share is far smaller than the sum of the individual single-mediator shares (approximately 50%, 50%, and 18%, which together would imply roughly 118%), revealing substantial effect overlap across different pathways under the common underlying logic of reduced information asymmetry. Shen et al. [13] and Yan et al. [35] focused on the single pathway of industrial upgrading, whereas this paper simultaneously incorporates the financial channel and factor allocation channel into the analysis, finding both to be equally important, which corroborates the conclusions of Nagaraj [27] and Lee et al. [26]. The heterogeneity analysis further reveals that the effect differences in the human capital dimension far exceed those in the city scale dimension, indicating that data interpretation capacity is more critical than economic aggregates. Although Cai and Ma [6] and Einav and Levin [7] discussed the non-rivalry characteristics of data factors, they did not adequately discuss dependence on complementary conditions. This paper supplements key empirical facts in this regard. At the regional level, the non-monotonic pattern of strong eastern, moderate western, and collapsing central challenges the naive expectation that effects decrease linearly with the level of economic development. Western cities, by receiving injections of institutional resources such as transfer payments and counterpart assistance, exhibit a pattern of weak endowments combined with strong policy yielding relatively high effects, providing new mechanistic annotations for the coordinated development research. In summary, this paper achieves the separation analysis of "green production process" and "environmental terminal outcome" in IGG dimensional decomposition, reveals the effect overlap characteristics and unified underlying logic in mechanism identification, and discovers the steep gradient of complementary condition constraints and non-monotonic regional patterns in heterogeneity analysis, adding new empirical facts to factor endowment theory in the era of data factors.

Limitations, Risks and Future Research

The first measurement limitation concerns the treatment variable. Following the prevailing literature, public data openness is operationalized as a binary indicator equal to one once a city has launched a public data open platform. This definition captures the extensive margin of data openness but does not account for qualitative differences across platforms, such as the number of datasets released and the availability of machine-readable formats.

Formats and application programming interfaces (APIs), data update frequency, or the breadth of thematic categories covered. Future research could construct a continuous openness index—drawing on metrics used by the World Wide Web Foundation or the European Open Data Maturity assessment—and re-estimate the treatment effects to test whether the magnitude of IGG gains increases with platform depth and quality. Second, the present analysis focuses on the positive effects of public data openness, but a balanced assessment must also recognize potential risks. Widespread data release can raise privacy and data-security concerns, especially when datasets contain granular information on individuals or firms. Open platforms may also increase exposure to cyber-attacks and widen inter-city digital divides if less-developed cities lack the analytical capacity to use newly released data. Moreover, because data flows tend to be concentrated in economically advanced regions, open-data policies could reinforce rather than reduce spatial inequalities unless complementary capacity-building measures are in place. Future work should therefore evaluate the trade-off between transparency and privacy, design differentiated data-security protocols, and examine whether open-data governance frameworks can mitigate unequal digital permeability across cities. Third, the empirical analysis treats the platform launch as a staggered adoption event and estimates average treatment effects over the sample period. Future work could exploit richer event-study and instrumental-variable designs, as well as spatial econometric specifications, to further identify dynamic effects and cross-city spillovers of public data openness.

9. Conclusions and Policy Recommendations

Against the dual backdrop of deepening digital economy and green and low-carbon transformation, how data factors empower IGG is an unavoidable question for China's high-quality development. Public data openness provides new insights for decoding this question through the logic of institutional openness. Based on panel data of 295 prefecture-level cities in China from 2010 to 2024, this paper systematically examines the impact effects, mechanisms, and heterogeneity characteristics of public data openness on urban IGG using the DDML method, drawing the following main conclusions: Public data openness has a significant positive effect on urban IGG, with positive improvements observed across the dimensions of economic development, social equity, green production and consumption, and environmental protection. This conclusion holds under multiple robustness checks. Mechanism analysis further reveals that data openness operates through three pathways: first, driving the penetration of digital inclusive finance and broadening the service coverage of green finance; second, mitigating factor market distortions and guiding factors to flow toward green and efficient sectors; and third, enhancing information accessibility and accelerating green technology diffusion and environmental information dissemination. However, this promoting effect is not uniformly distributed but is highly dependent on cities' own "absorptive capacity." Cities with larger scales, more abundant human capital, and higher initial IGG levels benefit more. From a regional perspective, the effect exhibits a pattern of "strong east—moderate west—weak central," reflecting that the depth of data dividend release depends on the maturity of digital infrastructure and market mechanisms.

Based on the above research conclusions, this paper proposes the following policy recommendations.

9.1. Driving the Transition of Data Openness from "Quantity" to "Quality" Through Institutional Standards

Although data open platforms have been initially established in various localities, the prevalent problem of "prioritizing launch over maintenance" leads to lagging data updates, inconsistent formats, and missing APIs, resulting in open data that is "visible but unusable." It is recommended to establish classification and grading standards and quality assessment mechanisms for public data openness at the national level, prioritizing the implementation of high-value datasets directly related to green growth, such as environmental monitoring, corporate carbon emissions, and green credit approval, and to incorporate data openness quality into local government performance evaluation, using institutionalization to drive openness effectiveness.

9.2. Connecting the Service Loop Between Data Openness and Green Finance

Research indicates that digital inclusive finance is the core transmission pathway through which data openness drives inclusive green growth. In practice, green financing for SMEs still faces the "information silo" dilemma—banks struggle to obtain enterprises' environmental compliance and energy consumption data, resulting in high green credit approval costs and low coverage. It is recommended to promote direct data interface connections between public data platforms and financial institutions, embedding corporate environmental credit, carbon emissions, and emission permits into the green credit approval process, reducing banks' due diligence costs, and ensuring that data dividends effectively benefit the green financing needs of SMEs and underdeveloped regions.

9.3. Leveraging Data Transparency to Optimize Factor Allocation

The mechanism through which data openness mitigates factor market distortion indicates that information asymmetry is an important root cause of factor misallocation. Currently, the allocation of key factors such as land transfer, capacity quotas, and carbon emission allowances in some cities still lacks open and transparent query channels, which easily breeds rent-seeking and inefficient allocation. It is recommended to embed mandatory information disclosure clauses in factor market-oriented allocation reforms, promoting the full-chain disclosure of allocation rules, processes, and results for the above factors, using data visibility to constrain allocation behavior and guide factors to flow in an orderly manner toward green and efficient sectors.

9.4. Implementing a Spatially Balanced Strategy Combining Differentiated Advancement with Cross-Regional Coordination

Heterogeneity tests and regional gradient analysis jointly indicate that the depth of data openness dividend release depends on cities' "absorptive capacity," and that eastern spillover effects can benefit the central and western regions. Therefore, for eastern megacities, the focus should be on opening high-granularity, real-time data to release deep innovation dividends. For central and western small and medium cities, priority should be given to addressing digital infrastructure shortcomings and providing supporting data literacy training to prevent data openness from becoming an "impractical concept." On this basis, leveraging national strategic platforms such as the coordinated development of the Beijing–Tianjin–Hebei region, the integration of the Yangtze River Delta, and the Guangdong–Hong Kong–Macao Greater Bay Area, cross-regional sharing mechanisms for environmental monitoring, green technology patents, and industrial chain carbon emission data should be pioneeringly established, using data interconnection to promote the

diffusion and spillover of green technology from the east to the central and western regions, gradually narrowing the growth gap of "strong east—weak central—moderate west.

Author Contributions: Conceptualization, W.L.; methodology, Z.Y. and W.L.; software, L.Z.; validation, Z.Y. and L.Z.; formal analysis, L.Z.; investigation, W.L.; resources, W.L.; data curation, Z.Y. and W.L.; writing—original draft preparation, Z.Y. and L.Z.; writing—review and editing, Z.L.; visualization, L.Z.; supervision, W.L.; project administration, W.L. and Z.L.; funding acquisition, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

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

References

1. Wang, J.; Zhu, W.W. Mechanisms and Strategies of Data Factors Empowering Chinese-Style Modernization during the 15th Five-Year Plan Period. *Econ. Issues* **2026**, 1–8.
2. World Bank. *Inclusive Green Growth: The Pathway to Sustainable Development*; World Bank Publications: Washington, DC, USA, 2012.
3. Barbier, E.B. The Green Economy Post Rio+20. *Science* **2012**, *338*, 887–888. [CrossRef]
4. Zhang, T.; Li, J. Network Infrastructure, Inclusive Green Growth, and Regional Disparities: Causal Inference Based on Double Machine Learning. *J. Quant. Technol. Econ.* **2023**, *40*, 113–135.
5. Yu, H.; Li, J. The Impact of Network Infrastructure Construction on Inclusive Green Growth: Empirical Evidence from the "Broadband China" Strategy. *Ind. Technol. Econ.* **2024**, *43*, 129–138.
6. Cai, Y.; Ma, W. The Impact of Data Factors on High-Quality Development and the Constraints of Data Flow. *J. Quant. Technol. Econ.* **2021**, *38*, 64–83.
7. Einav, L.; Levin, J. Economics in the Age of Big Data. *Science* **2014**, *346*, 1243089. [CrossRef]
8. Gu, K.; Sun, H. Measurement and Spatiotemporal Evolution Analysis of China's Urban Inclusive Green Growth Index. *Stat. Decis.* **2025**, *41*, 86–91.
9. Wang, Q.; Wang, Q.M. Can Data Factor Aggregation Drive Inclusive Green Growth? A Comparative Analysis of Tangible and Intangible Digital Capital. *Mod. Manag. Sci.* **2025**, 52–65.
10. Dong, Z.; Yin, C.; Zhang, L.; Zhang, Z.; Cui, H. Towards Inclusive Green Growth: Synergistic Effects of Digital Infrastructure and Energy Transition Policies in China. *J. Environ. Manag.* **2025**, *395*, 127993. [CrossRef]
11. Zhang, X.; Wan, G.; Zhang, J.; He, Z. Digital Economy, Inclusive Finance, and Inclusive Growth. *Econ. Res. J.* **2019**, *54*, 71–86.
12. Ai, M.; Wang, P.; Bu, Y. Climate Policy and Inclusive Green Growth: The Role of China's Low-Carbon City Pilot Policy. *J. Clean. Prod.* **2025**, *519*, 145959. [CrossRef]
13. Shen, N.; Zhou, J.; Zhang, G.; Wu, L.; Zhang, L. How Does Data Factor Marketization Influence Urban Carbon Emission Efficiency? A New Method Based on Double Machine Learning. *Sustain. Cities Soc.* **2025**, *119*, 106107. [CrossRef]
14. Dong, L.; Zhu, X.; Yang, L.; Jiang, G. Unleashing the Power of Data Element Markets: Driving Urban Green Growth Through Marketization, Innovation, and Digital Finance. *Int. Rev. Econ. Financ.* **2025**, *99*, 104070. [CrossRef]
15. Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Inf. Syst. Manag.* **2012**, *29*, 258–268. [CrossRef]
16. Zeleti, F.A.; Ojo, A.; Curry, E. Exploring the Economic Value of Open Government Data. *Gov. Inf. Q.* **2016**, *33*, 535–551. [CrossRef]
17. Lnenicka, M.; Nikiforova, A.; Clarinval, A.; Luterek, M.; Rudmark, D.; Neumaier, S.; Kević, K.; Rodríguez Bolívar, M.P. Sustainable Open Data Ecosystems in Smart Cities: A Platform Theory-Based Analysis of 19 European Cities. *Cities* **2024**, *148*, 104851. [CrossRef]
18. Finck, M.; Müller, M.-S. Access to Data for Environmental Purposes: Setting the Scene and Evaluating Recent Changes in EU Data Law. *J. Environ. Law* **2023**, *35*, 109–131. [CrossRef]
19. Sun, B. The Impact of Public Data Openness on Urban Green Innovation: A Quasi-Natural Experiment Based on Government Public Data Open Platforms. *J. Zaozhuang Univ.* **2025**, *42*, 89–98.
20. Lei, H.; Li, Q.; Chen, H. How Public Data Openness Drives Urban Innovation and Entrepreneurship. *Wuhan Financ.* **2024**, 70–78.

21. Ji, Y.; Wang, Y.N.; Wang, S.; Wang, R. Public Data Openness and Urban Entrepreneurial Activity: Evidence from Business Registration Data. *World Econ. Pap.* **2025**, 36–51.
22. Cai, Y.; Sun, M. Data Factor Sharing and Urban Entrepreneurial Vitality: Empirical Evidence from Public Data Openness. *J. Quant. Technol. Econ.* **2024**, *41*, 5–25.
23. Zhan, S.; Wang, Y.Z. How Public Data Openness Enhances Urban Innovation Capacity. *Mod. Manag. Sci.* **2025**, 94–105.
24. Fei, W.; Liang, X. The Synergistic Effect of Public Data Open Platforms and Data Trading Platform Openness on Enterprise Innovation Boundaries. *J. Macro-Qual. Res.* **2026**, *14*, 13–27.
25. Mao, Y.; Hu, N.; Leng, T.; Liu, Y. Digital Economy, Innovation, and Firm Value: Evidence from China. *Pac.-Basin Financ. J.* **2024**, *85*, 102355. [CrossRef]
26. Lee, C.C.; Wang, C.W.; Purnama, M.Y.I.; Sharma, S.S. Digitalization and Firms' Debt Maturity: Do Financial Constraints and Uncertainty Matter? *Pac.-Basin Financ. J.* **2024**, *85*, 102399. [CrossRef]
27. Nagaraj, A. The Private Impact of Public Data: Landsat Satellite Maps Increased Gold Discoveries and Encouraged Entry. *Manag. Sci.* **2022**, *68*, 564–582. [CrossRef]
28. Chen, Q.; Ding, X.; Wang, T. Can Open Government Data Improve Urban Innovation? Empirical Research from a Dual Perspective in China. *Gov. Inf. Q.* **2025**, *42*, 102059. [CrossRef]
29. Jiang, H.; Sun, S.L.; Jiang, P.C. Public Data Openness and Urban Economic Resilience: Empirical Evidence from the Launch of Government Public Data Platforms. *Econ. Surv.* **2025**, 94–106.
30. Qiao, Z.; Teng, M.R.; Dai, M.Y. Can Public Data Resource Sharing Enhance Supply Chain Resilience? A Quasi-Natural Experiment Based on Public Data Open Platforms. *World Econ. Stud.* **2025**, 88–104.
31. Huang, D.; Yang, F.; Zheng, J. Government Public Data Openness and the Coordinated Development of Urban Digitalization and Greenization: Evidence from the Launch of Public Data Open Platforms. *China Popul. Resour. Environ.* **2026**, *36*, 114–126.
32. Du, M.; Zhang, S. The Synergistic Effect of Public Data Openness on Pollution Reduction and Carbon Reduction: Quasi-Natural Experimental Evidence from the Launch of Government Data Platforms. *J. Henan Norm. Univ. (Philos. Soc. Sci.)* **2026**, *53*, 77–85.
33. Acemoglu, D.; Aghion, P.; Bursztyn, L.; Hemous, D. The Environment and Directed Technical Change. *Am. Econ. Rev.* **2012**, *102*, 131–166. [CrossRef]
34. Liu, X.; Qin, C.; Liu, B.; Ahmed, A.D.; Ding, C.J.; Huang, Y. The Economic and Environmental Dividends of the Digital Development Strategy: Evidence from Chinese Cities. *J. Clean. Prod.* **2024**, *440*, 140398. [CrossRef]
35. Yan, H.; Zhao, Z.; Zheng, Y.; Huang, J. Evaluating the Energy-Saving and Efficiency-Enhancing Potential of Data Factor Marketization: Empirical Evidence from 270 Cities in China. *J. Environ. Manag.* **2025**, *378*, 124686. [CrossRef]
36. Wang, F.; Li, H.H. The Common Prosperity Effect of Public Data Openness: Empirical Evidence from Macro and Micro Perspectives. *Mod. Econ. Res.* **2026**, 15–32.
37. 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]
38. 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]
39. Safarov, I.; Meijer, A.; Grimmelikhuijsen, S. Utilization of Open Government Data: A Systematic Literature Review of Types, Conditions, Effects and Users. *Inf. Polity* **2017**, *22*, 1–24. [CrossRef]
40. Hughes-Cromwick, E.; Coronado, J. The Value of US Government Data to US Business Decisions. *J. Econ. Perspect.* **2019**, *33*, 131–146. [CrossRef]
41. Fudan University Laboratory for Digital and Mobile Governance. *China Local Government Data Openness Report (First Half of 2025)*; Fudan University: Shanghai, China, 2025.
42. He, J.F.; Zhou, Y.L. Data Factor Marketization, New Quality Productive Forces, and Urban–Rural Common Prosperity: A Quasi-Natural Experiment Based on the Establishment of Data Trading Platforms. *Inq. Into Econ. Issues* **2025**, 91–109.
43. Zhang, M.; Xia, Z. The Impact of Public Data Openness on Basic Public Service Level and Its Mechanism: An Integrated Perspective of Attention Guidance, Economic Empowerment, and Social Inclusion. *Urban Probl.* **2025**, *362*, 29–41.
44. Liu, R.; Lin, S.; He, C. The Mechanism and Verification of Digital Inclusive Finance Promoting Green Inclusive Growth. *Ecol. Econ.* **2024**, *40*, 55–62.
45. Zhang, Y.; Tan, P. The Impact of Government Data Openness on Urban Green Total Factor Productivity. *Financ. Account. Mon.* **2025**, *46*, 123–128.
46. Zhao, C.; Cheng, G.; Zhang, J. Digital Economy Empowering Urban Inclusive Growth: From the Perspective of Effective Factor Markets. *Enterp. Econ.* **2026**, *45*, 113–124.
47. Zhou, X.; Wu, W. Measurement and Analysis of China's Inclusive Green Growth. *J. Quant. Technol. Econ.* **2018**, *35*, 3–20.
48. Han, X.; Li, J. Digital Intelligence Empowering Urban Inclusive Green Growth. *China Popul. Resour. Environ.* **2025**, *35*, 177–189.

49. Jiang, T. Mediation Effects and Moderation Effects in Empirical Research on Causal Inference. *China Ind. Econ.* **2022**, 100–120.
50. Sun, C.; Lin, Y. The Impact of Factor Market Distortion on Marine Economic Efficiency in China's Coastal Areas. *Mar. Sci. Bull.* **2021**, *40*, 369–378.

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.