Sustainability in Environment
ISSN 2470-637X (Print) ISSN 2470-6388 (Online)
Vol. 11, No. 4, 2026
www.scholink.org/ojs/index.php/se

Original Paper

Research on the Impact of Data Element Agglomeration on Urban Energy Efficiency

Wang Jian[1]

[1] College of Economics and Management, Hunan University of Technology, Zhuzhou, 412007, China

Received: March 27, 2026 Accepted: September 09, 2026 Online Published: September 18, 2026
doi:10.22158/se.v12n4p1 URL: http://dx.doi.org/10.22158/se.v12n4p1

Abstract

Data elements, as a new type of production factor, inject new momentum into enhancing urban energy efficiency and hold significant importance for achieving China's energy conservation, emission reduction, and green development goals. Based on panel data from 280 prefecture-level cities in China from 2011 to 2023, this study empirically examines the relationship between data element agglomeration and urban energy efficiency using a multi-period difference-in-differences model. The findings reveal that data element agglomeration significantly enhances urban energy efficiency. Mechanism tests indicate that data element agglomeration improves urban energy efficiency by promoting energy-saving technological progress and optimizing energy allocation. Heterogeneity analysis shows that the effect of data element agglomeration on urban energy efficiency is more pronounced in cities where officials have long-term performance perspectives, environmental regulations are stringent, and in resource-based cities. Further analysis demonstrates that data element agglomeration has a positive spatial spillover effect on urban energy efficiency, with the optimal spillover range observed within 200KM to 650KM from the city. This study provides important insights for governments to improve data element institutions and formulate energy conservation and emission reduction policies.

Keywords

Data Element Agglomeration, Urban Energy Efficiency, Energy-Saving Technological Progress, Energy Allocation Efficiency, Spatial Spillover Effect

1. Introduction

Against the backdrop of frequent fluctuations in the global energy market and complex and volatile geopolitics, as a major energy-importing country, China faces severe challenges to its energy security. The stability and security of energy supply are vital to China's overall economic development and social stability. In recent years, sharp fluctuations in international crude oil and natural gas prices have directly driven up supply costs in the domestic energy market. At the same time, geopolitical conflicts may obstruct energy transportation routes, seriously threatening the security of China's energy supply chain. By contrast, in China's energy consumption market, with sustained economic development and rising living standards, energy demand has increased year by year. In 2024, China's total energy consumption reached 5.96 billion tons of standard coal equivalent, a year-on-year increase of 4.3%. However, China's per capita energy resource endowment is relatively low, and per capita reserves of key fossil fuels such as coal, oil, and natural gas are all below the global average. The contradiction between resource constraints and growing demand has become increasingly prominent, placing enormous pressure on energy supply. Therefore, China urgently needs to improve energy efficiency to address energy security issues. However, existing energy management systems and models are still in the initial stage of digitalization and suffer from problems such as asymmetric energy supply and demand information, insufficient coordination, and delayed responses, making it difficult for them to provide strong support for the efficient operation of urban energy.

In this context, to resolve the dilemma of energy transition, policies such as the "Data Elements ×" Three-Year Action Plan (2024-2026) have been introduced, aiming to unleash the synergistic value of data elements and clarify the path for driving their deeper development. As electronic data resources that have been processed and analyzed and possess economic and social value, data elements, with characteristics such as non-rivalry and replicability, have become a key engine for driving urban energy efficiency upgrading. They can be simultaneously used by multiple entities, effectively solving problems such as asymmetric energy supply and demand information and insufficient coordination. On the one hand, through real-time collection and transmission, data elements break down information barriers in energy production, transmission, and consumption, achieving a dynamic balance between supply and demand. On the other hand, by integrating data from multiple energy systems and building an integrated dispatch model, they directly optimize energy conversion and transmission efficiency, facilitating efficient urban energy utilization.

To this end, this paper uses the establishment of national-level big data comprehensive pilot zones as a dummy variable for data element agglomeration and empirically examines its relationship with urban energy efficiency. The possible marginal contributions are as follows: First, it focuses on the deep coupling between data elements and energy efficiency and, by building a data-driven energy technology innovation system, releases the multiplier value of data elements in the energy field. Second, from the two perspectives of energy-saving technological progress and energy allocation efficiency, it deeply explores the enabling paths through which data element agglomeration empowers urban energy efficiency, comprehensively explains the underlying logic, and provides practical guidance for improving urban energy efficiency. Third, from the government perspective, through subgroup analysis based on officials' views on political performance and the intensity of government environmental regulation, it reveals the internal mechanism of government in the relationship between data elements and energy efficiency, helping it improve assessment mechanisms. Fourth, with the help of spatial effect analysis, it breaks through the assumption of independent observations, captures spatial correlations, distinguishes local measures from global impacts, and provides a decision-making basis for regional coordinated development and cross-regional governance.

2. Policy Background and Theoretical Analysis

2.1 Policy Background

With the rapid development of information technology, big data has become a core national strategic resource. In August 2015, the State Council issued the Action Outline for Promoting the Development of Big Data. The following month, Guizhou Province took the lead in building the first comprehensive pilot zone. In October 2016, the second batch list was announced, including the following types: cross-provincial coordinated development models such as the Beijing-Tianjin-Hebei region and the Pearl River Delta; local pioneering demonstration models such as Shanghai and Henan; and infrastructure coordinated promotion models such as Inner Mongolia. Relying on regional pilots, these pilot zones aim to innovate big data development paradigms and practical paths, deepen the integration and open sharing of data resources, promote technological innovation and industrial linkage, and provide replicable practical experience and scalable models for deepening big data applications nationwide, thereby fully unlocking the potential of data elements. The Development Report of the National Big Data (Guizhou) Comprehensive Pilot Zone 2023 shows that Guizhou Province has ranked among the regions with the largest scale and strongest capacity of domestically produced intelligent computing resources, and its data centers are accelerating their transition toward intelligent-computing priority and integrated storage and computing. It can thus be seen that the establishment of big data comprehensive pilot zones has effectively promoted the regional agglomeration of data elements, creating conditions for studying the impact of data element agglomeration on urban energy efficiency.

2.2 Theoretical Analysis

2.2.1 The Impact of Data Element Agglomeration on Urban Energy Efficiency

Energy efficiency measures the degree of matching between "energy input" and "effective output," that is, obtaining as much effective output that is genuinely meaningful to the target as possible with as little energy input as possible. Therefore, this paper theoretically analyzes the impact of data element agglomeration on urban energy efficiency mainly from the two perspectives of energy input and energy output.

In terms of energy input, data element agglomeration can effectively reduce energy consumption while maintaining a given output. On the one hand, the agglomeration and sharing of data elements within an urban area generate significant network effects, thereby raising their marginal output level. Under the condition of an increased marginal output of data elements, firms, based on the principle of revenue maximization, tend to partially substitute data elements for energy inputs under a given output target, thereby achieving a more efficient mode of production organization. This substitution effect reduces the share of energy in total inputs and improves energy efficiency[1]. On the other hand, capital and energy, as well as labor and energy, generally exhibit low elasticity of substitution and are complementary; increases in capital or labor input are often accompanied by increases in energy consumption[2]. However, the introduction of data elements improves the marginal output of capital and labor by optimizing capital allocation efficiency and enhancing labor productivity, thereby strengthening the ability of capital to substitute for energy and of labor to substitute for energy, shifting the relationship between them from complementarity to substitution[3]. Ultimately, in energy-intensive enterprises, capital and labor inputs can be reduced, thereby lowering energy dependence.

In terms of energy output, data element agglomeration can increase energy output given a fixed energy input. On the one hand, by virtue of their non-rivalrous and non-depletable characteristics, data elements can promote the sharing and diffusion of knowledge about energy technologies at extremely low marginal cost, thereby reducing information acquisition costs and trial-and-error costs in the process of technological R&D. This process, by increasing the social knowledge stock and improving the efficiency of knowledge reallocation, accelerates the innovation process of energy technologies and expands their application scope[4]. The result is that, without increasing energy input, it pushes outward the production possibility frontier of energy conversion, ultimately achieving an increase in marginal energy output. On the other hand, when producers find it difficult to accurately predict the dynamics of energy supply and demand, they tend to adopt conservative decision-making strategies, causing the actual output of energy input to fall below its potential efficiency level. The introduction of data elements, by integrating dispersed information throughout the entire production process, provides more accurate decision support for the coordinated allocation of energy with capital, labor, and other factors, reducing the uncertainty caused by demand decision bias and factor mix imbalance, thereby reducing ineffective energy consumption caused by information lag or decision errors. Ultimately, the actual output of energy input can more fully approach its potential output level. In summary, this paper proposes the following hypothesis:

H1: Data element agglomeration can improve urban energy efficiency.

2.2.2 The Enabling Paths through Which Data Element Agglomeration Empowers Urban Energy Efficiency

Within the theoretical framework of biased technological progress, the price effect and the market size effect are the main factors determining the direction of technological progress[5]. From the perspective of the price effect, the marginal cost of data elements approaches zero, and their unit use cost declines as the degree of agglomeration increases. Meanwhile, influenced by factors such as geopolitical conflicts, energy prices show an upward trend, causing the relative price gap between the two types of factors to gradually narrow. According to Hicks' induced technological innovation theory, when the relative price of a certain production factor decreases, it induces technological progress to be biased toward saving the relatively expensive factor. Therefore, guided by changes in the relative prices of data elements and energy factors, R&D resources will concentrate in the field of energy-saving technologies, making innovation entities more inclined to develop energy-saving technological paths that substitute data elements for energy factors[6]. From the perspective of the market size effect, the high concentration of data in cities forms a data market with network effects, reducing the marginal cost of data-driven technological innovation and, through the knowledge spillover effect, increasing the probability of success in technological innovation, thereby raising firms' expected returns. Against this background, developing technological solutions that can substitute abundant data elements for scarce energy factors has higher economic value. As price takers in the market, firms will accordingly adjust their technology choice strategies and allocate more R&D resources to the direction of energy-saving technological innovation[7]. Based on this, this paper proposes the following hypothesis:

H2: Data element agglomeration empowers the improvement of urban energy efficiency through energy-saving technological progress.

Data element agglomeration can promote the optimal allocation of energy factors, thereby improving urban energy efficiency. From the perspective of energy misallocation, energy allocation is often constrained by information asymmetry and coordination costs, resulting in spatial misallocation. Due to their mobility and non-rivalrous characteristics, data elements can break the segmented state of energy markets, form an energy market system, and reduce the degree of information asymmetry[8]. In a unified factor market, energy price signals can more truthfully reflect differences in energy scarcity across regions, thereby guiding energy from low marginal output sectors to high marginal output sectors, reducing the degree of energy misallocation, and ultimately improving overall energy efficiency[9]. From the perspective of transaction costs, the accumulation and interaction of data elements drive their value to grow exponentially, accelerating the process by which market information becomes more complete. When market participants can make decisions based on sufficient information, information search costs decrease significantly. At the same time, increased price transparency weakens imbalances in bargaining power, prompting more reasonable transaction terms, thereby reducing the likelihood of moral hazard and opportunistic behavior and lessening the need for contract performance monitoring. As transaction costs fall to a sufficiently low level, market governance mechanisms gradually replace hierarchical governance and become the dominant form[10]. Driven by this mechanism, the agglomeration of data elements promotes the transformation of energy allocation from centralized to distributed, making energy allocation more closely aligned with actual demand scenarios, thereby improving energy utilization efficiency. Based on this, this paper proposes the following hypothesis:

H3: Data element agglomeration empowers the improvement of urban energy efficiency by optimizing energy allocation.

2.3 Model Specification and Variable Description

2.3.1 Model Specification

To test Hypothesis H1, that is, to identify the causal relationship between the establishment of national-level big data comprehensive pilot zones and urban energy efficiency, this paper constructs a difference-in-differences model (1) for empirical testing based on the differences in the timing of prefecture-level cities' inclusion in the national big data comprehensive pilot zones. For Hypotheses H2 and H3, energy-saving technological progress and energy allocation efficiency are respectively introduced to construct a mechanism testing model (2).

$$Gtfpe_{it} = \partial_0 + \partial_1 Policy_{it} + \partial_2 X_{it} + V_i + V_t + \varepsilon_{it} \tag{1}$$

$$M_{it} = \partial_0 + \partial_1 Policy_{it} + \partial_2 X_{it} + V_i + V_t + \varepsilon_{it} \tag{2}$$

Where i and t denote city and year, respectively; $Gtfpe_{it}$ represents the energy efficiency of city i in year t; M_{it} represents energy-saving technological progress and energy allocation efficiency; $Policy_{it}$ indicates whether city i was designated as a national-level big data comprehensive pilot zone in year t. X_{it} denotes control variables that may affect the estimation results; V_i and V_t represent city fixed effects and year fixed effects, respectively; and ε_{it} denotes the random error term.

2.3.2 Variable Description

(1) Dependent Variable

The dependent variable is energy efficiency (Gtfpe). This paper takes labor, capital, and energy as inputs, regional GDP as the desired output, and industrial sulfur dioxide, industrial smoke and dust, and industrial wastewater emissions as undesirable outputs, and uses the SBM-Malmquist-Luenberger index method to measure the green total factor energy efficiency of each prefecture-level city. Compared with energy consumption per unit of regional GDP (ec/gdp), the advantages of green total factor energy efficiency are as follows: First, it considers urban heterogeneity. Energy consumption per unit of regional GDP adopts a unified energy consumption standard and ignores development differences among cities, which easily leads to distortion in energy efficiency evaluation. In contrast, green total factor energy efficiency can be measured based on the characteristics of each city, avoiding one-size-fits-all bias. Second, it provides more comprehensive indicators. Energy consumption per unit of regional GDP considers only two aspects, namely energy consumption and economic output, whereas actual energy efficiency is jointly affected by multiple inputs such as capital and labor, as well as desired outputs and undesirable outputs. The green total factor model incorporates multidimensional variables, more comprehensively reflecting the factors affecting energy efficiency and making the measurement results more accurate and reliable.

(2) Explanatory Variable

According to the approval documents of the National Development and Reform Commission, the Ministry of Industry and Information Technology, and the Cyberspace Administration of China, China launched the construction of national big data comprehensive pilot zones in eight regions, including the Beijing-Tianjin-Hebei region. The establishment of the pilot zones, as a strong policy signal, clarifies to the market the policy orientation that these regions will carry out pioneering trials in the field of data institutional innovation. This clear institutional expectation provides stable development plans for various market entities and can effectively guide data elements and related resources to agglomerate in the pilot zones. In view of this, this paper selects the "national-level big data comprehensive pilot zone" as a dummy variable for data element agglomeration. If a city is approved as a pilot zone in a certain year, then the city is assigned a value of 1 in that year and subsequent years, and 0 in other years.

(3) Mechanism Variables

Energy-saving technological progress (tec) refers to a form of technological progress that can improve the utilization efficiency of energy factors relative to other production factors. Under the condition of keeping the output level unchanged, it helps reduce the relative demand for energy. This paper adopts a nested CES production function and estimates it based on the normalized supply-side system method. The specification of the two-level nested CES production function is shown in Equation (3), and the normalized equations are shown in Equations (4) to (7), respectively. On the basis of estimating factor-augmenting technological progress, the level of energy-saving technological progress is further calculated according to Equation (8).

$$Y_t = \left\{ \partial (A_{Lt} L_t)^{\varepsilon} + (1-\partial)\left[(1-\beta)(A_{Et}E_t)^{\varpi} + \beta (A_{Kt}K_t)^{\varpi}\right]^{\frac{\varepsilon}{\varpi}} \right\}^{\frac{1}{\varepsilon}} \tag{3}$$

$$\log \frac{\rho_{E_t}}{\rho_{E_0}} = \varepsilon \log \xi - \varepsilon \log \frac{Y_t}{Y_0} + \varepsilon \log\left(e^{\frac{\gamma_l}{\lambda_l}(t^{\lambda_l}-1)} \frac{L_t}{L_0}\right) \tag{4}$$

$$\begin{aligned} \log \frac{\rho_{L_t}}{\rho_{L_0}} &= \varepsilon \log \xi - \varepsilon \log \frac{Y_t}{Y_0} + \varpi \log\left(e^{\frac{\gamma_e}{\lambda_e}(t^{\lambda_k}-1)} \frac{E_t}{E_0}\right) \\ &+ \frac{\varepsilon-\varpi}{\varpi} \log\left[\xi\left(e^{\frac{\gamma_e}{\lambda_e}(t^{\lambda_e}-1)} \frac{E_t}{E_0}\right)^{\varpi} + (1-\beta)\left(e^{\frac{\gamma_k}{\lambda_k}(t^{\lambda_k}-1)} \frac{K_t}{K_0}\right)^{\varpi}\right] \end{aligned} \tag{5}$$

$$\begin{aligned} \log \frac{\rho_{K_t}}{\rho_{K_0}} &= \varepsilon \log \xi - \varepsilon \log \frac{Y_t}{Y_0} + \varpi \log\left(e^{\frac{\gamma_k}{\lambda_k}(t^{\lambda_k}-1)} \frac{K_t}{K_0}\right) \\ &+ \frac{\varepsilon-\varpi}{\varpi} \log\left[\xi\left(e^{\frac{\gamma_e}{\lambda_e}(t^{\lambda_e}-1)} \frac{E_t}{E_0}\right)^{\varpi} + (1-\beta)\left(e^{\frac{\gamma_k}{\lambda_k}(t^{\lambda_k}-1)} \frac{K_t}{K_0}\right)^{\varpi}\right] \end{aligned} \tag{6}$$

$$\begin{aligned} \log \frac{Y_t}{Y_0} &= \log \xi + \frac{1}{\varepsilon} \log \Bigg\{ (1-\partial)^{1-\varepsilon} \rho_{L_0} \left(e^{\frac{\gamma_l}{\lambda_l}(t^{\lambda_l}-1)} \frac{L_t}{L_0}\right)^{\varepsilon} + \partial^{1-\varepsilon} (1-\rho_{L_0})^{1-\frac{\varepsilon}{\varpi}} \\ &\left[(1-\beta)^{1-\varpi} \rho_{E_0} \left(e^{\frac{\gamma_e}{\lambda_e}(t^{\lambda_e}-1)} \frac{E_t}{E_0}\right)^{\varpi} + \xi^{1-\varpi} \rho_{K_0} \left(e^{\frac{\gamma_k}{\lambda_k}(t^{\lambda_k}-1)} \frac{K_t}{K_0}\right)^{\varpi}\right]^{\frac{\varepsilon}{\varpi}} \Bigg\} \end{aligned} \tag{7}$$

$$tec = \varpi d\left(\frac{A_{Kt}}{A_{Et}}\right) / \frac{A_{Kt}}{A_{Et}} dt \tag{8}$$

where A_{Lt}、A_{Kt}、A_{Et} denote labor-augmenting technological progress, capital-augmenting technological progress, and energy-augmenting technological progress, respectively; α is the share parameter of the labor factor; β is the share parameter of capital in the energy-capital composite; ε is the elasticity of substitution between labor and the energy-capital composite; ϖ is the elasticity of substitution between energy and capital; t is time; ζ is the scale factor of the equation; ρ_L、ρ_K、ρ_E are the output shares of the three factors; ρ_{L0}、ρ_{K0}、ρ_{E0} are the arithmetic means of the output shares of each factor; and Y_0、L_0、K_0、E_0 are the average values of output, labor, capital, and energy consumption, respectively. tec is energy-saving technological progress; when tec > 0, it indicates that technological progress is biased toward energy saving; when tec < 0, it represents technological progress biased toward capital saving.

Energy allocation efficiency (eae) refers to a city's ability to maximize output under a given production technology and management level by optimizing the combination, allocation, and utilization of energy and other production factors. This paper extends the traditional production function into a three-factor Cobb-Douglas production function with energy, capital, and labor, and obtains it by taking logarithms. Next, based on the estimated coefficients, the marginal output of the energy factor is calculated as $\alpha Y_{it}/E_{it}$. Then, the degree of deviation is obtained according to the ratio of the marginal output of the energy factor to the energy price. Finally, the relative ratio of each city's degree of deviation to the maximum degree of deviation among cities in that year is used as the measurement indicator of urban energy allocation efficiency.

$$\ln Y_{it} = C + \partial \ln E_{it} + \beta \ln K_{it} + \gamma \ln L_{it} + \varepsilon_{it} \tag{9}$$

where Y_{it}、E_{it}、K_{it} 和 L_{it} denote gross regional product, energy consumption, capital stock, and number of employees, respectively; α, β, and γ denote the output elasticity of energy, the output elasticity of capital, and the output elasticity of labor, respectively; C is the constant term; and ε is the random error term.

(4) Control Variables

To objectively identify the causal effect between national-level big data comprehensive pilot zones and urban energy efficiency, and to avoid result bias caused by omitted variables, this paper, with reference to previous studies, selects the following control variables: population density (den), measured by the ratio of population size to urban area; industrial structure (stru), measured by the ratio of value added of the secondary industry to GDP; industrial structure (ind), measured by the ratio of value added of industry to GDP; human capital level (hc), measured by the ratio of the number of students enrolled in regular higher education institutions to the total population at year-end; wage level (wage), represented by the average wage of employees.

2.3.3 Data Sources and Descriptive Statistics

Based on the principle of data availability, this paper selects 280 prefecture-level and above cities in China from 2011 to 2023 as the research sample. Among them, there are 80 cities in the treatment group and 200 cities in the control group. The data mainly come from the China City Statistical Yearbook, the list of national-level big data comprehensive pilot zones, and statistical yearbooks and statistical bulletins of various provinces and cities. In addition, to address missing values in the original data, this paper uses linear interpolation to fill them in, ultimately obtaining the empirical data. The descriptive statistics of each variable are shown in Table 1.

Table1. Descriptive Statistics of Variables

Variable Type	Variable Name	Sample Size	Mean	Standard Deviation	Minimum	Maximum
Dependent Variable	Urban Energy Efficiency	3640	0.3332	0.1067	0.1026	0.9964
Explanatory Variable	Data Element Agglomeration	3640	0.1679	0.3738	0	1
Mechanism Variables	Energy-Saving Technological Progress	3640	-0.9689	16.7884	-59.2826	69.3669
	Energy Allocation Efficiency	3640	0.3223	0.1615	0.0033	0.9915
Control Variables	Population Density	3640	5.7342	0.9911	1.7360	9.0886
	Industrial Structure	3640	0.4475	0.1099	0.0873	0.8205
	Industrial Sector Structure	3640	0.3807	0.1366	0.0276	2.8487
	Human Capital Level	3640	0.0200	0.0262	0.0001	0.4976
	Wage Level	3640	63702.53	24075.56	10216	320626

3. Empirical Results Analysis

3.1 Baseline Regression Results

Table 2 reports the impact of data element agglomeration on urban energy efficiency. First, the results in Column (1) show that, without adding control variables and without fixing city and year fixed effects, the core explanatory variable is significantly positive, initially verifying that data element agglomeration helps improve urban energy efficiency. Second, after further adding control variables and fixed effects, the regression coefficient of the core explanatory variable in Column (4) is 0.0288 and significant, indicating that data element agglomeration can improve urban energy efficiency by about 2.88%. Hypothesis H1 is thus confirmed.

Table 2. Baseline Regression Results

Variable	Gtfpe			
	(1)	(2)	(3)	(4)
Policy	0.0568***	0.0249***	0.0228***	0.0288***

	(0.0064)	(0.0064)	(0.0064)	(0.0063)
Control Variables	No	No	Yes	Yes
City Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes
Observations	3640	3640	3640	3640
R^2	0.023	0.688	0.126	0.7

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

3.2 Parallel Trend Test and Dynamic Effect Analysis

The application of a multi-period difference-in-differences model requires satisfying the prerequisite assumption of parallel trends. This assumption holds that before the establishment of the pilot zones, there is no significant difference in energy efficiency between pilot and non-pilot cities, and efficiency begins to improve only after establishment. To verify this assumption, this paper uses an event study method for testing, and the results are shown in Figure 1. Before the establishment of national-level big data comprehensive pilot zones, the coefficients are all insignificant, and there is no obvious difference in energy efficiency between the two types of cities; after establishment, urban energy efficiency improves, satisfying the parallel trend test.

From the perspective of dynamic effects, the big data comprehensive pilot zones become significant only three years after establishment, indicating that policy implementation has a certain time lag, possibly because a large amount of digital infrastructure needs to be built in the initial stage of establishment to meet the needs of big data processing and application. From the overall trend analysis, the policy effect of national-level big data comprehensive pilot zones is significantly positive and shows an upward trend, implying that its role in improving urban energy efficiency will continue to strengthen in the future.

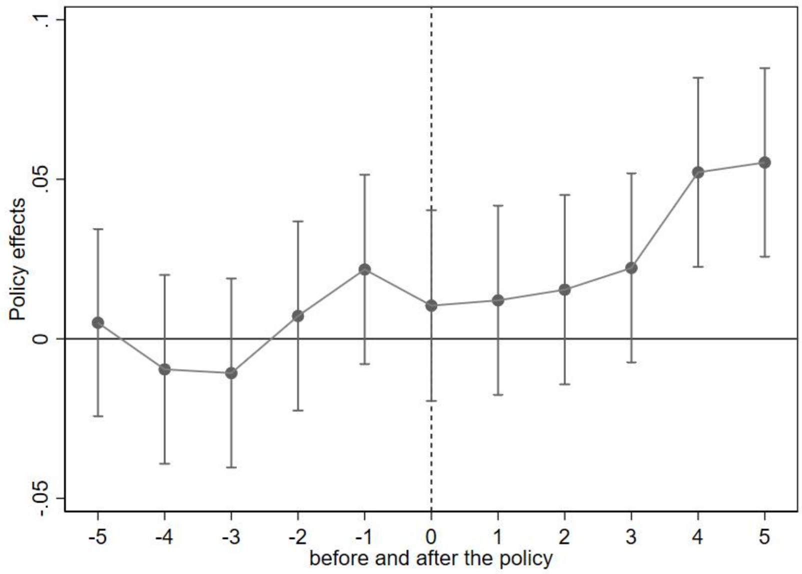

Figure 1. Parallel Trend Test

3.3 Endogeneity Test

Since the approval of national-level big data comprehensive pilot zones is not conducted with cities as the unit, it is difficult to fully account for all observable or unobservable factors, and endogeneity problems may exist. Therefore, this paper selects the shortest distance between the administrative centers of each prefecture-level city and each node city in the "Eight Vertical and Eight Horizontal" optical cable backbone transmission network, and multiplies it by the one-period lagged number of national internet users to construct an instrumental variable. In terms of relevance, prefecture-level cities close to the node cities of the "Eight Vertical and Eight Horizontal" optical cable backbone network have geographical advantages in data element agglomeration. In terms of exogeneity, the node cities of the backbone network had been basically completed by 1998, and the distances between each prefecture-level city and them belong to the category of historical geography, being fixed and stable, making it difficult to affect current data element agglomeration. Therefore, the instrumental variable satisfies the relevance and exogeneity assumptions. The test results are shown in Table 3. The coefficient of the instrumental variable is significant, and it passes the underidentification test and the weak instrument test; the coefficient of the core explanatory variable in the second stage is significantly positive, indicating that after excluding the influence of endogeneity, the driving effect of data element agglomeration on the improvement of urban energy efficiency remains robust.

Table 3. Endogeneity Test Results

Variable	First Stage	Second Stage
	Policy	Gtfpe
iv	-2.58e-06*** (3.00e-07)	
Policy		0.4016*** (0.0573)
Control Variables	Yes	Yes
City Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Kleibergen-Paap rk	79.20	
LM statistic	[0.00]	
Kleibergen-Paap rk	73.94	
Wald F statistic	{16.38}	

Note. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses, p-values in square brackets, and critical values at the 10% level for the Stock-Yogo weak identification test in curly braces.

3.4 Robustness Tests

3.4.1 PSM-DID Test

When selecting big data comprehensive pilot zones, the state may give priority to certain cities as pilots due to factors such as economic level, research capacity, geographical location, and infrastructure, causing the estimation results to be affected by self-selection bias. To address the problem of non-random selection, this paper uses propensity score matching to construct a control group highly similar in characteristics to the treatment group, and conducts a difference-in-differences model test on the matched sample. Specifically, control variables are first selected as covariates, and propensity scores are calculated using 1:1 nearest-neighbor matching to screen out the control group most similar to the treatment group. Next, the difference-in-differences model is used again to re-estimate the impact of data element agglomeration on urban energy efficiency. The results in Column (1) of Table 4 show that the regression coefficient of the matched sample is still significantly positive, further verifying the robustness of the conclusion that data element agglomeration is conducive to improving urban energy efficiency.

3.4.2 Placebo Test

Even if control variables are included in the baseline regression and city and time effects are fixed, it is still difficult to eliminate bias caused by unobservable factors. Therefore, this paper uses a placebo test, randomly drawing from the sample the same number of cities as the actual number of cities where national big data comprehensive pilot zones were established in each year as the treatment group, with the rest as the control group, re-estimating, and repeating the experiment 500 times. The results are shown in Figure 2. After randomization, the estimated coefficients are mostly distributed near zero and approximately normal, and their corresponding P-values are generally higher than 0.1, significantly different from the coefficients of the baseline model, indicating that the interference of unobservable factors on the core conclusion is limited, effectively verifying the robustness of the research results.

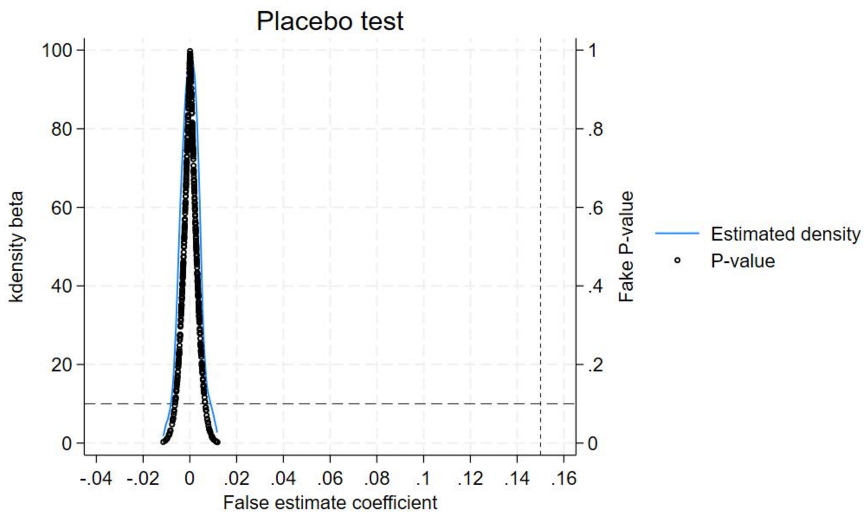

Figure 2. Placebo Test

3.4.3 Excluding Interference from Other Policies

Considering that the "Broadband China" and "Smart City" initiatives issued during the same period played a role in big data agglomeration to a certain extent, this paper incorporates them into the regression model to control for their impact on the regression results. As shown in Column (2) of Table 4, the coefficient of data element agglomeration remains significantly positive, indicating that after excluding interference from other policies, its positive impact on urban energy efficiency remains unchanged.

3.4.4 Double Machine Learning

To control for unpredictable factors affecting regional energy efficiency, avoid the "curse of dimensionality" and estimation bias caused by control variables, and also considering the possible nonlinear relationships among variables, this paper adopts a double machine learning model for robustness testing. As shown in Column (3) of Table 4, the coefficient of data element agglomeration is still positive and significant, verifying the robustness of the regression results.

3.4.5 Excluding Municipalities Directly Under the Central Government and Provincial Capital Cities

Given that municipalities directly under the Central Government and provincial capital cities differ significantly from ordinary prefecture-level cities in economic growth, digitalization level, and energy efficiency, this paper excludes municipalities directly under the Central Government and provincial capital cities for regression testing to eliminate interference caused by differences in administrative levels. As shown in Column (4) of Table 4, the coefficient of data element agglomeration is still positive and significant, indicating that after removing cities with higher administrative levels, the research results remain robust.

Table 4. Robustness Test Results

Variable	(1)	(2)		(3)	(4)
	PSM-DID	Controlling concurrent policies Broadband China	Smart City	Double Machine Learning	Excluding Municipalities and Provincial Capitals
Policy	0.0214** (0.0093)	0.0291*** (0.0063)	0.0281*** (0.0063)	0.0269*** (0.0079)	0.0227*** (0.00644)
Policy1		0.0053 (0.00597)			
Policy2			-0.0118 (0.00763)		
Control Variables	Yes	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	1360	3640	3640	3640	3263
R^2	0.757	0.7	0.7	/	0.699

Note. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses.

3.5 Mechanism Tests

3.5.1 Analysis of the Enabling Mechanism of Energy-Saving Technological Progress

As shown in Column (1) of Table 5, the coefficient of data element agglomeration is significantly positive, indicating that it can promote energy-saving technological progress, thereby improving urban energy efficiency. Data element agglomeration breaks down the barriers of information asymmetry, reduces uncertainty in the process of technological innovation activities, and enables market entities to more accurately identify key nodes of energy loss and potential optimization spaces, thereby guiding

R&D resources to be allocated to energy-saving technology fields with higher returns. This guiding effect is not only reflected in the optimization of existing technological routes but also further promotes the integration and reconstruction of cross-domain knowledge.

3.5.2 Analysis of the Enabling Mechanism of Energy Allocation Efficiency

The coefficient in Column (2) of Table 5 is positive and significant, indicating that data element agglomeration can optimize energy allocation. In the energy sector, data isolation once existed among entities such as power generation enterprises, grid companies, energy-using enterprises, and government regulatory departments, and the non-rivalrous nature of data elements breaks down this barrier. Energy-using enterprises share data on production energy use and equipment energy efficiency with government regulatory departments, based on which the government formulates scientific energy policies and energy-saving standards to guide enterprises in energy conservation and emission reduction; enterprises, in turn, adjust their energy use structure and production processes according to policies to improve energy utilization. In addition, data elements achieve agglomeration through algorithmic mining and cross-domain integration, eliminating information silos, transforming dispersed energy potential into synergistic efficiency, and ultimately achieving a significant improvement in energy efficiency, attaining a "1+1 > N" effect. Hypothesis H2 is confirmed.

Table 5. Mechanism Test Results

Variable	(1)	(2)
	tec	eae
Policy	3.5413***	0.0185***
	(1.1849)	(0.0069)
Control Variables	Yes	Yes
City Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Observations	3640	3640
R^2	0.35	0.74

Note. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses.

3.6 Heterogeneity Analysis

3.6.1 Heterogeneity Analysis of Officials' Views on Political Performance

Different views on political performance among local government officials may exert differentiated impacts on the construction of big data comprehensive pilot zones. Since the construction of these pilot zones is a systematic project with a long cycle and involving multidimensional changes, officials with a long-term performance orientation may be inclined to implement national policies and actively promote data element agglomeration; by contrast, officials with a short-term performance orientation may pursue short-term performance projects for rapid promotion. Based on the above analysis, this paper expects that in cities with officials holding a long-term performance orientation, data element agglomeration is more likely to improve urban energy efficiency. Under the normal assessment mechanism, it is common to be promoted to mayor around the age of 50, but under the trend of younger cadres, officials in this age group have limited promotion opportunities and are more likely to form a long-term view of political performance. Therefore, this paper groups cities by the age of the incumbent mayor at the time, classifying those aged 50 or above as likely possessing a long-term performance orientation and those under 50 as likely possessing a short-term performance orientation. As shown in Columns (1) and (2) of Table 6, in cities with officials holding a long-term performance orientation, the effect of data element agglomeration on improving urban energy efficiency is more significant, consistent with the above expectation.

3.6.2 Heterogeneity Analysis of Environmental Regulation Intensity

Differences in government environmental regulation intensity may produce different impacts on urban energy efficiency. The traditional regulatory model has efficiency bottlenecks, forcing cities to improve energy efficiency through management optimization, and data element agglomeration provides a "testing ground" for this. Cities with strong environmental regulation pay more attention to the role of data elements in optimizing environmental management and can better leverage data element agglomeration to improve energy efficiency and the environment; cities with weak environmental regulation may be content with the status quo and fail to effectively take the "free ride" of data element agglomeration to improve energy efficiency. Based on the above analysis, this paper expects that in cities with strong environmental regulation, the effect of data element agglomeration on improving energy efficiency is more significant. Accordingly, this paper constructs a measure of environmental regulation intensity based on the frequency of words related to "environmental protection" in government work reports, and divides cities into a "strong regulation group" and a "weak regulation group" according to the relative level of each city's value compared with the overall mean during the sample period. As shown in Columns (3) and (4) of Table 6, in cities with strong environmental regulation, the effect of data element agglomeration on improving urban energy efficiency is more significant, consistent with the above expectation.

3.6.3 Heterogeneity Analysis of Resource Endowment

Energy efficiency is greatly affected by regional resource endowments. Most resource-based cities, due to long-term reliance on traditional resource development models, generally suffer from the "resource curse," the essence of which is a lack of capacity for efficient resource integration, leading to low energy efficiency. Data element agglomeration can break this dilemma and improve energy efficiency by creating new production factors, reconstructing production relations, and optimizing resource allocation; in non-resource-based cities, energy consumption is mainly concentrated in buildings, transportation, and daily life, with a low proportion in high-energy-consuming sectors such as industry, leaving limited room for direct transformation by data elements. Based on the above analysis, this paper expects that in resource-based cities, the effect of data element agglomeration on improving energy efficiency is more significant. Accordingly, this paper refers to the Notice of the State Council on Issuing the National Sustainable Development Plan for Resource-Based Cities (2013–2020) to distinguish sample cities into resource-based cities and non-resource-based cities. As shown in Columns (5) and (6) of Table 6, in resource-based cities, the effect of data element agglomeration on improving urban energy efficiency is more significant, consistent with the above expectation.

Table 6. Heterogeneity Analysis Results

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Long-term Performance Orientation	Short-term Performance Orientation	Strong Regulation Intensity	Weak Regulation Intensity	Resource-based Cities	Non-resource-based Cities
Policy	0.0376*** (0.0095)	0.0089 (0.0084)	0.0347*** (0.0088)	0.0206** (0.0088)	0.0342*** (0.0087)	0.0216** (0.0086)
Chow Test Empirical P-value	0.0000		0.0000		0.0000	
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1651	1989	1586	2054	1482	2158
R^2	0.70	0.71	0.67	0.72	0.73	0.67

Note. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses.

4. Further Analysis: Spatial Effect Analysis

First, data elements possess the characteristic of non-rivalry. Data can be infinitely replicated and shared, with a usage cost approaching zero, and can radiate to surrounding areas at low cost, establishing a spillover foundation of "one party invests, multiple parties benefit." Second, data elements have strong penetrability and cross-boundary integration capability. Data can deeply penetrate fields such as the economy, government affairs, and people's livelihood, and their spillover effects diffuse in a network pattern, breaking industrial and regional boundaries. Finally, data elements have knowledge attributes and innovation-empowering value. The process of data element agglomeration is accompanied by knowledge production, achieving knowledge spillovers through talent mobility and technological cooperation. Based on the cross-regional impact characteristics of data element agglomeration, this paper uses spatial econometric models for verification, which not only helps cities improve energy efficiency through data element agglomeration but also promotes coordinated development among cities.

4.1 Spatial Autocorrelation Analysis

Before conducting spatial econometric analysis, it is necessary to calculate the Moran's I index to test the spatial correlation of data element agglomeration empowering the improvement of urban energy efficiency. The results are shown in Table 7. From 2011 to 2023, the Moran's I index of urban energy efficiency is significantly positive, indicating the existence of positive spatial correlation.

Table 7. Moran's I Results from 2011 to 2023

Year	Moran' s I	Z-statistic	P-value
2011	0.059	12.177	0.00
2012	0.035	7.566	0.00
2013	0.026	5.887	0.00
2014	0.030	6.719	0.00
2015	0.040	8.542	0.00
2016	0.025	5.627	0.00
2017	0.043	9.212	0.00
2018	0.042	8.966	0.00
2019	0.051	10.588	0.00
2020	0.047	9.881	0.00
2021	0.039	8.358	0.00
2022	0.039	8.347	0.00
2023	0.038	8.123	0.00

4.2 Spatial Spillover Effect Analysis

Before the regression, following the testing path proposed by Elhorst (2014)[11], this paper sequentially conducts the LM test, the fixed effects test for the SDM model, the Hausman test, and the simplification test of the SDM model. Taking the spatial Durbin model into comprehensive consideration, this paper constructs a spatial Durbin difference-in-differences model based on Equation

(1):

$$Gtfpe_{it} = \rho WGtfpe_{it} + \beta_1 Policy_{it} + \beta_2 X_{it} + \theta_1 WPolicy_{it} + \theta_2 WX_{it} + V_i + V_t + \varepsilon_{it} \quad (10)$$

where ρ is the spatial autocorrelation coefficient, W denotes the geographical distance matrix, and the meanings of the remaining coefficients are consistent with those in the previous text.

The results are shown in Table 8. The direct and indirect effects of data element agglomeration on the improvement of urban energy efficiency are both significantly positive, verifying that it has a spatial spillover effect and can influence surrounding cities. The reasons are as follows: First, the demonstration effect. The successful experiences of cities with data element agglomeration in energy management are easily learned from by other cities, promoting regional collaboration. Second, knowledge spillover. Data element agglomeration is accompanied by the application of advanced technologies, which spread to surrounding areas through channels such as enterprise cooperation, talent mobility, and academic exchange. Third, the platform economy effect. The large digital platforms on which data element agglomeration relies can optimize the matching of energy supply and demand through cross-regional data integration, thereby achieving efficiency improvement.

Table 8. Analysis of Spatial Spillover Effect Results

Variable	Gtfpe
Policy	0.0179**
	(0.0087)
W*Policy	0.116**
	(0.0533)
ρ	0.327***
	(0.1220)
Direct effect	0.0186**
	(0.0089)
Indirect effect	0.181**
	(0.0791)
Total effect	0.199***
	(0.0741)
Matrix type	Geographical matrix
Control Variables	Yes
City Fixed Effects	Yes
Year Fixed Effects	Yes
Observations	3640
R^2	0.037

Note. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses.

4.3 Analysis of Spatial Distance Attenuation Effect

In fact, constrained by factors such as administrative barriers, the radiation range of infrastructure, and technological adaptability, the spatial spillover effect of data element agglomeration often has geographical scope limitations. To this end, this paper will further verify whether there is a spatial distance attenuation effect on urban energy efficiency. This paper divides the geographical distance matrix by setting thresholds at intervals of every 50 km, up to 1000 km.

Table 9. Spatial Distance Attenuation Effect

Distance	Spatial Spillover Effect Coefficient	P-value	Distance	Spatial Spillover Effect Coefficient	P-value
50KM	0.076	0.17	550KM	0.341	0.00
100KM	0.034	0.58	600KM	0.219	0.00
150KM	0.063	0.37	650KM	0.130	0.05
200KM	0.082	0.27	700KM	0.030	0.61
250KM	0.146	0.08	750KM	-0.037	0.49
300KM	0.225	0.01	800KM	-0.016	0.73
350KM	0.419	0.00	850KM	-0.047	0.30
400KM	0.492	0.00	900KM	-0.034	0.44
450KM	0.527	0.00	950KM	-0.057	0.18
500KM	0.435	0.00	1000KM	-0.018	0.66

From Table 9, it can be seen that the spatial spillover effect of data element agglomeration on urban energy efficiency can be divided into three stages. The first stage is 0–200 km. In this interval, the spatial spillover effect coefficient is not significant. The main reasons are as follows: First, short-distance siphon lock-in. Data elements are highly concentrated in hub cities, forming a data highland, and neighboring cities find it difficult to absorb technological spillovers; instead, they experience a Matthew effect due to the siphoning of talent and capital. Second, competition for energy resources. The high energy consumption characteristics of data center clusters lead nearby cities to compete for energy supply and environmental capacity, squeezing surrounding energy allocation, and the negative effects dominate. Third, local institutional inertia. The energy policies and standards of nearby cities are more similar, reform resistance is concentrated, and the spatial spillover effect of data element agglomeration is weakened.

The second stage is 200–650 km. In this interval, the spatial spillover effect coefficient is significantly positive, representing an appropriate zone that is "neither neighboring competition nor distant irrelevance." On the one hand, the synergy effect outweighs the competition effect. Compared with the resource competition and homogenization problems among nearby cities, in this interval data transmission costs are controllable, and the information gap and technological gap form a driving force for gradient diffusion, promoting coordinated urban development. On the other hand, it is easier to form a vertical division of labor network. The industrial structure differences among medium-distance cities are relatively obvious, allowing the formation of a vertical division of labor model of "data element agglomeration for production—energy application and transformation," replacing homogeneous competition with complementary competition and amplifying the spatial spillover effect of data element agglomeration on energy efficiency.

The third stage is 650–1000 km. In this interval, the spatial spillover effect coefficient is not significant. The main reasons are as follows: First, the physical limitations of network infrastructure. Long-distance data transmission requires more computing power for encryption, compression, and error correction, and the diseconomies of scale in cities' returns lead to attenuation of the spillover effect. Second, data governance rules are not unified. Cities more than 650 km apart usually belong to different provincial-level administrative regions or economic blocks, and relevant energy consumption data are difficult to open to distant cities due to local protection policies, directly inhibiting the spillover effect. Third, industrial structure differences are significant. The industrial systems of distant cities differ greatly, lacking an "interface" to adapt to the needs of data elements, and the spillover path is blocked.

5. Research Conclusions and Policy Recommendations

5.1 Research Conclusions

Against the strategic background of global energy transition and climate change governance, how to improve urban energy efficiency has become a core issue for China in implementing the "dual carbon" goals, ensuring energy security, and promoting high-quality development. Based on the pilot data of big data comprehensive pilot zones in 280 prefecture-level cities in China from 2011 to 2023, the China City Statistical Yearbook, and other data, this paper empirically examines the impact and mechanism of data element agglomeration on urban energy efficiency through a multi-period difference-in-differences model. The research findings are as follows: (1) Data element agglomeration, presented in the form of the establishment of big data comprehensive pilot zones, can improve urban energy efficiency, and this conclusion still holds after robustness tests such as the PSM-DID test and double machine learning. (2) The mechanism test finds that data element agglomeration empowers the improvement of urban energy efficiency by promoting energy-saving technological progress and optimizing energy allocation. (3) The heterogeneity analysis finds that, from the government perspective, in cities with officials holding a long-term performance orientation and in cities with strong government environmental regulation, the effect of data element agglomeration on improving urban energy efficiency is stronger; from the perspective of city characteristics, resource-based cities are better able to use data element agglomeration to improve urban energy efficiency. (4) Further analysis finds that data element agglomeration has a positive spatial spillover effect on the improvement of urban energy efficiency, and the optimal spillover zone is between 200 km and 650 km.

5.2 Policy Recommendations

Based on the research conclusions of this paper, the following policy recommendations are proposed:

First, strengthen the institutional guarantee for data element agglomeration. In the energy field, it is necessary to improve the data element supply mechanism from two aspects: classified data rights confirmation and public data management. First, in the energy data dimension, implement a three-way classification of public, enterprise, and personal data, and establish an operating mechanism of separating the three rights of holding, processing, and operating according to source and generation characteristics. It should be clarified that energy enterprises, after lawfully collecting and processing non-sensitive data, enjoy the rights to hold, use, and receive returns, thereby incentivizing enterprises to participate in data supply. Second, for the massive public data generated in the course of government performance of duties, the collection, sharing, and open utilization should be strengthened. Following the principle of "raw data does not leave the domain, and data is available but not visible," value should be released to society through models, verification, and other forms. For data that does not contain personal information and does not impede public security, the scope of openness should be expanded; for data serving public governance and public welfare undertakings, conditional free use may be implemented; for data oriented toward industrial development, paid authorization may be explored. At the same time, it is strictly prohibited by law for undisclosed raw public data to directly enter the market, ensuring that the public interest is not eroded. In summary, improving the data element system can provide fundamental support for its efficient allocation and application in the urban energy field, break down institutional barriers to data elements at the top-level design level, and thereby systematically improve urban energy efficiency.

Second, use city characteristics as an anchor to precisely release the effectiveness of data elements. From the perspective of officials' views on political performance, on the one hand, establish an energy assessment system. By setting an indicator of "the contribution rate of data element application to energy efficiency improvement," require a certain contribution target to be met and include it as a bonus item in officials' promotion assessments; on the other hand, create data demonstration projects. The central government can provide special subsidies for data demonstration projects and publicize typical cases at national conferences to strengthen officials' competitive incentives. From the perspective of government environmental regulation, strengthen government regulation by building a territory-wide IoT monitoring network to collect energy consumption data in real time and automatically trigger warnings when energy consumption exceeds standards. At the same time, cities with weak regulation often have the problem of "loose and soft law enforcement," which needs to be solved by solidifying the evidence chain through data and strengthening credit punishment.

Third, leverage data element spillovers to achieve win-win regional energy efficiency. First, establish a cross-regional data element coordination mechanism. Based on the optimal spatial spillover radius of

200–650 km, plan and establish a number of "regional data-energy coordination circles." At the same time, core cities within urban agglomerations can take the lead in setting up regional data element cooperation funds to support joint research. Second, implement a digital infrastructure sinking project, in which central cities export mature data service models and technologies to surrounding cities, and at the same time dispatch technical teams to guide the digital infrastructure of surrounding cities, make up for their shortcomings, and amplify the spillover effect. Third, build "data element + industry" cross-city parks, and provide joint tax incentives to enterprises in these cross-city parks to encourage the deep integration of data elements and industrial entities.

References

[1] Cui P, Zhuang K. Digital economy and industrial structure transformation and income distribution[J]. Applied Economics, 2025, 57(35): 5215-5231.

[2] Green F, Stern N. China's changing economy: implications for its carbon dioxide emissions[J]. Climate policy, 2017, 17(4): 423-442.

[3] Alvarez-Cuadrado F,Van Long N,Poschke M.Capital–labor Substitution,Structural Change,and Growth[J].Theoretical Economics,2017,12(3):1229-1266.

[4] Cai Y Z, Ma W J. The impact of data elements on high-quality development and constraints on data flow[J]. Journal of Quantitative & Technological, 2021, 3(38): 64-84.

[5] Acemoglu D.Directed Technical Change[J].The Review of Economic Studies,2002,69(4):781-809.

[6] Acemoglu D,Aghion P,Bursztyn L,et al.The Environment and Directed Technical Change[J].American Economic Review,2012,102(1):131-166.

[7] Zhao X, Hu L, Chen X, et al. Energy-biased technological progress and green innovation: Evidence from Chinese cities[J]. Renewable Energy, 2025: 124079.

[8] Carbonara N, Nisar T M, Prabhakar G, et al. Open data business models and financial stability in open innovation environments[J]. R&D Management, 2025, 55(5): 1657-1670.

[9] Cohen F, Glachant M, Söderberg M. The impact of energy prices on energy efficiency: Evidence from the UK refrigerator market[R]. Grantham Research Institute on Climate Change and the Environment, 2015.

[10] Riordan M H,Williamson O E.Asset Specificity and Economic Organization[J].International Journal of Industrial Organization,1985,3(4):365-378.

[11] Elhorst J P.Matlab Software for Spatial Panels[J].International Regional Science Review,2014,37(3):389-405.