

Article

The Mechanism of Data Elements Application Empowering Urban–Rural Integrated Development

Jinyan Wen ^{1,*}, Xiaoyan Fan ¹ and Jialin Li ²¹ School of Marxism, Xi'an Jiaotong University, Xi'an 710049, China; fanxiaoyan1974@xjtu.edu.cn² School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an 710049, China; lijialin@xjtu.edu.cn

* Correspondence: 1037261962@stu.xjtu.edu.cn

Abstract

Efficient and unimpeded factor allocation constitutes the cornerstone of urban–rural integrated development. Using panel data from 281 prefecture-level cities in China spanning the period from 2011 to 2023, this study employs a two-way fixed-effects model to empirically investigate the direct effects and transmission channels of the application of data elements on urban–rural integrated development. The findings are as follows: (1) The application of data elements significantly promotes urban–rural integrated development. On average, a 1% increase in the level of data element application is associated with a 0.089% rise in the urban–rural integrated development index. This result remains robust across a battery of robustness tests, with pronounced regional heterogeneity. (2) Mechanism analysis demonstrates that data elements application drives urban–rural integrated development through three primary channels: facilitating labor mobility, mitigating capital misallocation, and elevating total factor productivity (TFP). (3) Heterogeneity analysis based on regional characteristics reveals that this promoting effect is more significant in regions with superior economic growth, more advanced industrial structures, and abundant human capital. (4) Spatial spillover analysis indicates that while data element application facilitates local urban–rural integrated development, it simultaneously exerts a negative spatial spillover effect on neighboring regions. Accordingly, policymakers should prioritize strengthening digital infrastructure in less developed regions and take the lead in establishing cross-jurisdictional data sharing platforms coupled with benefit compensation mechanisms, thereby fostering coordinated advancement and mutually beneficial outcomes in urban–rural integrated development across regions.

Keywords: data elements; urban–rural integrated development; labor mobility; capital misallocation; total factor productivity

Academic Editor: Wei Song

Received: 19 May 2026

Revised: 21 June 2026

Accepted: 24 June 2026

Published: 26 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Urban–rural integrated development stands as a core issue in the global modernization process and constitutes one of China's strategic objectives for achieving common prosperity and high-quality development. Historically, driven by urban-biased development strategies, traditional production factors such as capital, labor, and technology have continuously agglomerated in urban areas, whereas rural regions have been trapped in the dilemma of factor outflows and lagging development [1–4]. Consequently, the imbalance in urban–rural factor allocation has emerged as a primary obstacle to constraining

urban–rural integrated development. To dismantle the urban–rural dual barriers induced by the unidirectional flow of traditional production factors, there is an urgent need to explore novel driving mechanisms and pathways for factor empowerment. Data, as a critical production factor in the digital economy era, is profoundly reshaping conventional resource allocation patterns through its development, openness, circulation, and utilization [5]. In practical terms, the application of data elements has transformed virtual space into a new frontier for resource integration. This application not only enhances the allocative efficiency of capital, technology, and talent across urban and rural domains but also enables deep mining and iterative utilization of data value. By constructing novel infrastructure and factor support systems for urban–rural integrated development, data element application is fundamentally reshaping the overall landscape of urban–rural relations. Therefore, an in-depth analysis of the mechanistic pathways through which data elements affect urban–rural integrated development carries substantial theoretical value and practical significance for advancing coordinated and high-quality urban–rural development.

Since data factors were officially included in the category of production factors, research on data factors has achieved remarkable progress in both theoretical and empirical domains [6]. At the theoretical level, scholars argue that data elements are essentially derivatives of the combination of natural resources and labor inputs. They convert human activities and physical phenomena into digital traces that can be recorded, stored, and analyzed, thereby enabling the large-scale generation and utilization of data as production factors [7]. Derived from the transformation of knowledge and information, data elements possess distinctive characteristics, including strong permeability, non-excludability, and diminishing marginal costs [8]. At the empirical level, the facilitating effects of data elements application have attracted extensive scholarly attention. At the macro level, the application of data elements enables pattern identification and knowledge extraction from massive datasets. Through synergistic interactions with capital, technology, labor, and other production factors, it exerts a multiplier effect on TFP improvement [9,10]. At the micro level, data elements application helps reduce information asymmetry and transaction costs while fostering the emergence of new business models and innovative organizational forms. Particularly, data elements characterized by high value density and strong interconnectivity exhibit significant scale effects and network effects, which serve to strengthen firms' innovation capacity and market competitiveness [11]. Nevertheless, a strand of literature holds a cautious view, arguing that the large-scale deployment of data elements may give rise to data monopolies. By capitalizing on their data advantages, large platform enterprises can build up market entry barriers, which in turn aggravate factor allocation distortions and widen income inequality [12]. Moreover, the economic effects of data elements application are heterogeneous across industries, regions, and development stages. In regions with underdeveloped digital infrastructure, the expected benefits of data elements application may be undermined by the digital access divide [13].

The measurement and evaluation of urban–rural integrated development have undergone remarkable evolution. Early studies relied on single indicators, such as the urban–rural income ratio and consumption ratio [14,15]. While intuitive and data-accessible, these metrics failed to capture the multidimensional and complex nature of urban–rural integration. As research deepened, scholars progressively constructed comprehensive evaluation index systems encompassing multiple dimensions—including economic, social, demographic, and spatial integration—thereby providing a more holistic assessment of urban–rural integration levels [16]. Methodologically, measurement approaches have expanded from traditional statistical methods to incorporate emerging data sources and quantitative techniques, including the incorporation of nighttime light data to gauge economic activity intensity, the entropy method for objective weighting, and global principal component analysis for constructing composite indices [17]. Regarding the exploration of

influencing factors, existing literature has predominantly conducted analyses from traditional macro perspectives, such as industrial structure upgrading, infrastructure construction, fiscal transfer payments, and land institutional reforms [18]. In recent years, concurrent with the rise of the digital economy, the driving effects of digital economy dimensions—including digital inclusive finance, digital infrastructure, and digital technology applications—on urban–rural integration have increasingly emerged as a research focus. Owing to its inclusive characteristics for vulnerable groups, digital inclusive finance provides more accessible financial services for rural areas, raises the marginal returns on rural investment, facilitates the flow of capital to rural areas, and promotes the integration of urban and rural capital factors [19,20]. Digital infrastructure construction reshapes the urban–rural economic geography and dismantles obstacles to factor mobility, driving the gradual superposition and integration of agricultural and non-agricultural industries [21,22]. Digital technology effectively clears bottlenecks across the links of urban–rural production, circulation, distribution, and consumption, enhancing market synergy and accelerating the integration process [23]. Nevertheless, the aforementioned studies predominantly focus on the macro-level application manifestations of the digital economy, overlooking the essential attribute of data elements as the underlying resource of the digital economy. In fact, whether manifested as digital finance, digital infrastructure, or digital technologies, their empowerment of urban–rural integration is fundamentally an external manifestation of the cross-regional flow and value release of data elements. Therefore, it is necessary to uncover the intrinsic mechanisms through which data elements contribute to urban–rural integrated development from a micro-level perspective.

Overall, although the extant literature offers valuable insights into how data elements application promotes urban–rural integrated development, there are still some deficiencies. First, existing studies have not fully elucidated the mutually embedded and symbiotic relationship between data elements and traditional production factors. The theoretical logic through which data elements, via deep integration with traditional factors, unleash multiplier effects that reshape the urban–rural factor allocation landscape and propel integrated development requires further clarification. Second, there is a paucity of empirical research that incorporates data elements as a separate variable within the analytical framework of urban–rural integration. Most extant studies have subsumed data elements within proxy variables for the digital economy or digital technologies, without having constructed a dedicated measurement index system for the application level of data elements to identify their specific effects on urban–rural integrated development. Third, insufficient attention has been devoted to the mechanistic pathways and spatial effects through which data elements influence urban–rural integration. Through which channels do data elements affect urban–rural integrated development? Do these effects exhibit regional heterogeneity and spatial spillovers? These questions await more in-depth investigation.

Accordingly, the contributions of this paper are threefold. (1) It elucidates the mutually embedded and symbiotic relationship between data elements and traditional production factors, and clarifies the theoretical logic whereby data elements advance urban–rural integrated development via three transmission channels: facilitating labor mobility, mitigating capital misallocation, and elevating TFP. (2) It constructs an evaluation index system for the application of data elements from three interrelated dimensions—namely, available supply, unimpeded flow, and effective utilization—and employs nighttime light data to quantify the level of urban–rural integrated development, providing a methodological extension for identifying the effects of data factors on urban–rural integrated development. (3) Third, it examines the direct effects, transmission channels, regional heterogeneity, and spatial spillover effects of data elements application on urban–rural integrated development, thereby enriching the empirical evidence base on the developmental implications of data elements and providing a decision-making reference for the

implementation of differentiated policies to foster urban–rural integration. The remainder of this study is organized as follows. Section 2 develops the theoretical analysis and research hypotheses. Section 3 describes the model specification, data sources, and variable selection. Section 4 reports empirical results, including baseline regressions, heterogeneity analysis, and spatial effect analysis. Section 5 presents the mechanism analysis. The final section summarizes the main conclusions, discusses the policy implications, and highlights the limitations and future research directions.

2. Theoretical Analysis and Research Hypotheses

2.1. *The Direct Effect of Data Elements Application on Urban–Rural Integrated Development*

Factor allocation theory posits that production factors flow freely across sectors and regions in response to price signals until the marginal product of each factor converges, thereby achieving Pareto optimality in resource allocation [24]. However, the long-standing urban–rural dualistic structure in China has led to the institutional segmentation of factor markets—such as labor, capital, and land—imposing high transaction costs on cross-regional factor mobility and consequently engendering unbalanced urban–rural development. The intrinsic attributes of data elements—virtuality, strong permeability, and zero marginal cost—offer a novel pathway to overcome this dilemma, thereby facilitating urban–rural integrated development [25]. First, the virtuality of data elements facilitates the cross-boundary circulation of urban and rural information, thereby mitigating urban–rural information asymmetry. Data generated from rural economic activities can flow rapidly into urban markets, while urban capital, technology, and employment information can penetrate rural areas. This enables both the supply and demand sides to make decisions on a more informed basis, optimizing the allocation of production factors between urban and rural areas and promoting urban–rural integrated development. Second, the strong permeability of data elements application enables it to form a mutually embedded and symbiotic relationship with traditional production factors, driving urban–rural integrated development at higher levels and across broader domains. Specifically, land provides the spatial carrying capacity for the deployment of data elements applications, digital agriculture, and smart industrial parks; labor furnishes the contextual foundation for the application of data elements; capital supplies the material support for digital platforms; and technology offers the instrumental tools for data collection, processing, and analysis. In turn, data elements reciprocally enhance the allocative efficiency and marginal returns of traditional factors, rendering labor more skilled, capital more targeted, land more intensive, and technology more diffusive. Finally, the zero marginal cost characteristic of data elements amplifies the advantages of reuse-driven efficiency enhancement, substantially lowering barriers to information acquisition and knowledge dissemination. Network externality theory posits that each additional node in a network enhances the utility of existing nodes, and that network value grows non-linearly with user scale [26]. The urban–rural information networks and digital platforms constructed through data elements application similarly benefit from the continued expansion of urban–rural data nodes and platform users. As a result, the efficiency of data sharing and supply–demand matching is continuously improved, enabling urban and rural residents to share innovation resources on an equal footing and further advancing the equalization of public services and urban–rural integrated development. The theoretical analysis framework for the data elements application to enable urban–rural integrated development is shown in Figure 1. Accordingly, we propose the following hypothesis:

Hypothesis 1. *Data elements application exerts a significant positive effect on urban–rural integrated development.*

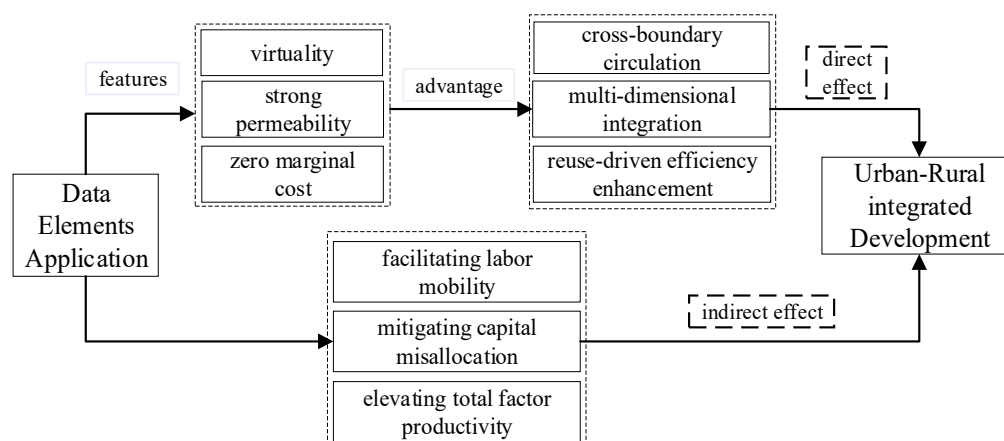


Figure 1. The theoretical analysis framework.

2.2. The Transmission Channels of Data Elements Application on Urban–Rural Integrated Development

The mutually embedded and symbiotic relationship between data elements and land factors is primarily manifested in the profound application of land data platforms, such as Geographic Information Systems (GIS) and digital cadastral systems. These platforms transform the current status, development potential, and spatial suitability of land use into computable digital representations, thereby facilitating the rational circulation and intensive utilization of urban–rural land resources. Transaction cost theory posits that the efficiency of market transactions is highly dependent on the costs of information search, contract negotiation, and enforcement monitoring, and that ill-defined property rights and information asymmetry can significantly elevate transaction costs while inhibiting resource allocation efficiency [27]. Digital cadastral systems, through unified spatial coordinates and ownership registration, clarify the boundaries of land property rights, reduce information search costs and title risks in land transactions, and render the ownership verification, planning coordination, and benefit distribution associated with the marketization of collectively owned operational construction land more transparent. Furthermore, GIS technology drives the transition of land resource allocation from an experience-dominated to a data-driven paradigm, enhancing the precision of spatial identification for urban–rural fringes, potential reclamation zones, and industrial undertaking areas. Given that land allocation between urban and rural areas in China is primarily realized through state institutional arrangements, use regulation, and policy-based adjustments, this study omits the discussion of land factors when examining the transmission channels through which data elements application affects urban–rural integrated development. The mobility and marketization degree of land factors remain considerably lower than those of labor and capital factors, rendering it difficult for land to generate significant mediating transmission effects driven by data elements in the short term. Instead, it focuses on production factors exhibiting stronger market attributes and higher mobility—specifically labor, capital, and technology—to elucidate the channels through which data elements application empowers urban–rural integrated development across three dimensions: facilitating labor mobility, mitigating capital misallocation, and elevating TFP.

2.2.1. Facilitating Labor Mobility

Labor, as the most fluid production factor, plays a pivotal role in advancing urban–rural integrated development through the optimal allocation between urban and rural areas. The effectiveness of labor mobility, however, is highly contingent upon information completeness. Information asymmetry theory posits that asymmetries between labor supply and demand—concerning skill profiles, job requirements, and mobility preferences—

can significantly elevate search costs and matching frictions, thereby inhibiting the orderly flow of labor [28]. The core value of data elements application resides in information transmission and the reduction of transaction costs [29], which facilitates the orderly flow of labor between urban and rural areas. Data elements application carries a vast amount of valid factual information, enabling urban centers to accurately gauge the quantity of labor supply, skill structures, and mobility intentions, thereby enhancing urban pull factors for rural labor [30]. Simultaneously, emerging industries underpinned by data elements—such as rural e-commerce, smart agriculture, and rural tourism—provide expanded development opportunities for individuals possessing relevant skills or entrepreneurial aspirations, attracting labor flows into rural areas [31].

Labor serves not only as the skill-bearing carrier for data elements application but also as the transmission agent for the bidirectional flow of information between urban and rural domains. Workers possessing digital competencies are better positioned to collect, interpret, and apply data with greater efficiency, thereby realizing the value inherent in data elements. Conversely, labor mobility between urban and rural areas objectively facilitates the cross-regional dissemination of digital literacy, technical expertise, and market intelligence. Consequently, the resulting two-way labor mobility contributes to accelerating urban–rural integrated development [32]. On the one hand, the migration of surplus labor to cities attenuates disguised unemployment in rural areas while providing low-cost labor for urban growth; this population agglomeration in cities raises the efficiency of resource utilization, generates scale effects, and promotes urban–rural integrated development. On the other hand, returning rural migrants deploy the skills, experience, and capital accumulated in cities to pursue development in their hometowns; university graduates and digitally skilled talents secure greater employment and entrepreneurial opportunities in new business formats, thereby fostering the diversification of rural industries and urban–rural integrated development (Figure 1). Accordingly, we propose the following hypothesis:

Hypothesis 2a. *Data elements application promotes urban–rural integrated development by facilitating urban–rural labor mobility.*

2.2.2. Mitigating Capital Misallocation

Capital serves as the material foundation for constructing data infrastructure, including cloud computing centers, communication networks, and Internet of Things terminals. Without sustained capital investment, the collection, storage, and computation of data elements would be impossible. In turn, data elements reshape the allocation logic of capital, shifting capital allocation from a traditional model reliant on collateral and subjective experience to a precision pricing model underpinned by data profiling and algorithmic models. Insufficient information acquisition and inaccurate risk pricing lead to capital misallocation [33]. Data elements application mitigates urban–rural capital misallocation by enhancing information transparency and risk assessment capabilities. Through the deep embedding of data elements, digital financial innovations—such as big data credit scoring and blockchain technology—construct more precise credit profiles for rural business entities, reduce their reliance on traditional collateral, and significantly expand their financing capacity. Concurrently, real-time production data generated by the agricultural Internet of Things and satellite remote sensing provide dynamic evidence for improving the precision of agricultural project risk assessments, guiding capital more effectively toward rural characteristic industries and small and medium-sized enterprises.

Lowering the degree of urban–rural capital misallocation and channeling capital toward sectors and regions with higher production efficiency contribute to urban–rural integrated development [34] in two respects. On the one hand, increased access to capital

enables rural areas to develop modern agriculture, rural service industries, and emerging non-agricultural sectors; this raises the value added of rural industries and gradually narrows the urban–rural gap in industrial structure and development levels. On the other hand, as capital returns gradually converge across urban and rural areas, capital is prompted to flow back in an orderly manner from over-concentrated cities to rural regions, thereby improving the equilibrium of capital allocation between urban and rural areas and substantially advancing urban–rural integrated development (Figure 1). Accordingly, we propose the following hypothesis:

Hypothesis 2b. *Data elements application promotes urban–rural integrated development by mitigating capital misallocation.*

2.2.3. Elevating Total Factor Productivity

The narrow concept of TFP, originating from neoclassical growth theory, primarily refers to technological progress. Technology diffusion theory further posits that the economic value of technological innovation depends not merely on the invention per se, but more critically on its diffusion speed and adoption breadth across disparate regions and industries [35]. Diffusion barriers tend to generate persistent technology gaps between urban and rural areas. By accelerating technology diffusion and knowledge spillovers through digital platforms and knowledge networks, data factor application shortens the transmission cycle of innovation outcomes from urban to rural areas, thereby comprehensively enhancing TFP across both agricultural and non-agricultural sectors [36,37]. In the agricultural sector, multi-source information carried by data elements—such as meteorological, soil, and crop growth data—provides scientific support for precision agriculture technologies, enabling the intelligent management of planting, irrigation, and fertilization processes, thereby significantly raising agricultural production efficiency. In the non-agricultural sector, data-driven technologies such as the industrial Internet and artificial intelligence promote the digital transformation of traditional industries, optimize production processes and resource allocation, and enhance the productive efficiency of manufacturing and services.

An increase in TFP implies higher output efficiency under given resource inputs, which is conducive to urban–rural integrated development [38] in two respects. On the one hand, the enhancement of agricultural production efficiency optimizes production processes, raises the quality and value added of agricultural output, provides cities with a high-quality agricultural product supply, and promotes the integrated development of urban and rural industries. On the other hand, technology diffusion and knowledge spillover effects accelerate the improvement of urban industrial efficiency and drive the technological upgrading and structural optimization of related rural industries, thereby realizing complementary and synergistic development between urban and rural industries. These dynamics generate a virtuous cycle encompassing factor mobility, industrial layout, and productivity enhancement, further advancing urban–rural integrated development (Figure 1). Therefore, we propose the following hypothesis:

Hypothesis 2c. *Data elements application promotes urban–rural integrated development by elevating TFP.*

3. Empirical Research Design

3.1. Econometric Model Specification

To empirically examine the impact of data factor application on urban–rural integrated development, we construct a two-way fixed-effects model as the baseline specification [39]. On the one hand, the inclusion of city fixed effects effectively controls for time-invariant city-specific heterogeneity that affects urban–rural integrated development,

such as geographic conditions, historical legacies, and resource endowments. On the other hand, the incorporation of time fixed effects captures time-varying shocks, including national macro digital economy policies, technological evolution, and macroeconomic fluctuations. This specification alleviates potential endogeneity concerns and ensures the consistency and robustness of parameter estimates. The model is specified as Equation (1):

$$URI_{it} = \alpha_0 + \alpha_1 DF_{it} + \sum_j \alpha_j Control_{it} + \eta_i + \nu_t + \varepsilon_{it} \quad (1)$$

where i and t denote the region and year, respectively; URI_{it} represents urban–rural integrated development in region i in year t ; DF_{it} denotes the level of data elements application in region i in year t ; and $Control_{it}$ is a vector of control variables. η_i represents region fixed effects, ν_t represents time fixed effects, and ε_{it} is the stochastic error term.

3.2. Variable Definitions

3.2.1. Explained Variable

The explained variable in this study is urban–rural integrated development (URI). We utilize the spatial distribution of nighttime light data to characterize urban–rural integrated development [40]. The intensity and distribution of nighttime light data reflect the vitality of the urban–rural spatial system and its structural changes, thus capturing the evolutionary characteristics of spatial connectivity and economic integration between urban and rural areas. To define the spatial boundaries of built-up areas, a statistical data verification method is used for spatial identification. The core principle is to identify the threshold value of nighttime light data such that the total area extracted based on this threshold minimizes the difference from the total urban built-up area reported in the *China City Statistical Yearbook* for the corresponding year, and simultaneously minimizes the sum of squared errors between the two. The specific procedure is as follows: The extreme values of nighttime light data DN are set as the iteration range, and the mean value is taken as the initial potential threshold DN_x . The area extracted at threshold DN_x is denoted as S_1 and compared with the total built-up area S_2 reported in the *China City Statistical Yearbook*. If $S_1 > S_2$, DN_x is assigned to the maximum value DN_{max} to contract the threshold interval; if $S_1 < S_2$, DN_x is assigned to the minimum value DN_{min} . After multiple iterations, the process terminates when S_1 and S_2 are approximately equal, and the sum of squared errors is minimized; the resulting DN_x represents the optimal threshold for extracting the built-up area. Using this threshold as the boundary, rasters where $DN > DN_x$ are classified as urban areas, while rasters where $DN < DN_x$ are classified as rural areas. Combining this classification with the weighted average method, the calculation formula for urban–rural integrated development is specified in Equation (2):

$$URI_{it} = \frac{\sum (DN_{it}^c \times Count_{it}^c)}{\sum Count_{it}^c} / \frac{\sum (DN_{it}^u \times Count_{it}^u)}{\sum Count_{it}^u} \quad (2)$$

In Equation (4), the value of URI_{it} ranges between 0 and 1. A value closer to 1 indicates a higher level of urban–rural integrated development in the region. The superscript c denotes rural areas, and the superscript u represents urban areas. DN_{it} represents the nighttime light intensity of region i in year t , and $Count_{it}$ represents the number of raster grid cells corresponding to the nighttime light intensity in region i in year t , which reflects the actual physical coverage in geographic space.

3.2.2. Explanatory Variable

The explanatory variable in this study is the data elements application (DF). Drawing on the market linkage approach, which is centered on the supply, flow, and utilization of data elements, this study constructs a comprehensive evaluation index system for city-level data element development across three dimensions: available supply, unimpeded flow, and effective utilization [41]. The entropy weight method is adopted for measurement (Table 1). In terms of the available supply dimension, we select the number of employees in the information transmission, computer services, and software industries to reflect the reserve of professional talents on the data elements supply side. Fiscal expenditure on science and technology is adopted to capture government policy support for data elements development. In terms of the unimpeded flow dimension, the number of subscribers to internet services is used to indicate the infrastructure conditions for domestic data elements circulation, while the number of international internet subscribers measures the scope of international data elements circulation. In terms of the effective utilization dimension, software business revenue is selected to gauge the capacity of data elements to be converted into actual economic outcomes. E-commerce transaction volume reflects the breadth of data elements application in the platform economy, capturing the enabling effect of data elements on business models.

Table 1. Comprehensive Evaluation Index System for Data Elements.

Criterion Layer	Indicator Layer	Specific Measurement	Unit	Data Sources
Available supply	Digital labor supply	Employees in the information transmission, computer services, and software	person	<i>China City Statistical Yearbook</i> ; EPS Data Platform
	Fiscal expenditure on science and technology	Expenditure for science/Local government general budget revenue	%	<i>China City Statistical Yearbook</i> ; EPS Data Platform
Unimpeded flow	Domestic network coverage	Number of subscribers of internet services	10,000 households	<i>China City Statistical Yearbook</i> ; EPS Data Platform
	International Network Interconnection	Number of international internet subscribers	10,000 households	<i>China Statistical Yearbook on Science and Technology</i> ; EPS Data Platform
Effective utilization	Software industry revenue	Software business income	10,000 yuan	<i>China Information Industry Yearbook</i> ; EPS Data Platform
	Platform transaction status	E-commerce transactions	100 million yuan	CNRDS Data Platform

3.2.3. Control Variables

The control variables include economic growth (EGR), industrial structure upgrading (ISU), infrastructure level (UIL), degree of openness (Open), fiscal expenditure pressure (FP), and education investment level (EIL). The definitions and measurement methods for the primary variables are summarized in Table 2. Regions experiencing faster economic growth typically possess stronger capacities for urban–rural coordination and exhibit narrower disparities in income and public services [42]. Industrial structure upgrading constitutes the industrial foundation for urban–rural integration, as it directly affects the quality of rural labor transfer and the urban–rural income gap through channels such as employment creation and wage growth [43]. Infrastructure integration represents an inherent dimension of urban–rural integration in its own right. Cities with a higher degree of openness typically feature a higher level of marketization and faster alignment with the global digital

economy, which shapes the progression of urban–rural integrated development [44]. Fiscal expenditure pressure serves as a critical institutional determinant, as it directly governs the intensity of local government investment in rural revitalization and the equalization of urban–rural public services [45]. Finally, education investment enhances the human capital stock across both urban and rural areas, narrows urban–rural gaps in educational attainment and capability endowment, and promotes urban–rural integrated development [46].

Table 2. Variable Definitions and Measurements (2011–2023).

	Variable Name	Measurement	Unit	Abbreviation	Data Sources
Explained Variable	Urban–rural integrated development	Based on nighttime light data	index (0–1)	URI	DMSP/OLS and NPP/VIRS nighttime light data
Explanatory Variable	Data elements	Based on the comprehensive evaluation index system	index (0–1)	DF	Detailed data sources are reported in Table 1
Control Variables	Economic growth	GRP growth rate	%	EGR	<i>China Urban and Rural Construction Statistical Yearbook</i> ; EPS Data Platform
	Industrial structure upgrading	Value-added of the tertiary industry/gross regional product	%	ISU	
	Infrastructure level	Road surface area per capita	sq.m	UIL	
	Degree of openness	Amount of foreign capital actually utilized/gross regional product	%	Open	
	Fiscal expenditure pressure	(Local government general budget expenditure–revenue)/revenue	%	FP	
	Education investment level	Expenditure for education/gross regional product	%	EIL	

3.3. Data Sources

This study utilizes a panel sample of 281 prefecture-level cities in China over the period 2011–2023. The data are primarily sourced from the *China Urban and Rural Construction Statistical Yearbook*, the *China City Statistical Yearbook*, the *China Statistical Yearbook on Science and Technology*, the *China Information Industry Yearbook*, statistical bulletins on national economic and social development, various municipal statistical yearbooks, as well as data platforms including the CNRDS and EPS database. For specific indicators not reported, which account for approximately 3.2% of total sample observations, missing values are imputed using trend extrapolation and linear interpolation methods. All continuous variables are tested using a two-tailed test at the 1% level to mitigate the impact of extreme outliers.

4. Empirical Results Analysis

4.1. Baseline Regression

Table 3 reports the estimation results of Model (1). Column (1) presents the regression results controlling exclusively for regional and year fixed effects. The estimation results indicate that the coefficient of the core explanatory variable is 0.088, which is statistically significant at the 5% level. Columns (2)–(7) display the regression results following the sequential inclusion of control variables. The estimated coefficients of the data elements application remain statistically significant and positive at the 5% level across these specifications. In the full model encompassing all control variables in Column (7), the regression coefficient for data elements application is 0.089, thereby supporting Hypothesis 1.

This suggests that, leveraging its characteristics of high permeability, reusability, and near-zero marginal costs, data elements application significantly facilitates the efficient flow of information, resources, and production factors between urban and rural areas, ultimately driving urban–rural integrated development.

Table 3. Baseline Regression Results.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	URI	URI	URI	URI	URI	URI	URI
DF	0.088 ** (2.269)	0.091 ** (2.353)	0.094 ** (2.444)	0.091 ** (2.370)	0.094 ** (2.440)	0.091 ** (2.357)	0.089 ** (2.301)
EGR		0.186 * (1.876)	0.269 *** (2.686)	0.266 *** (2.658)	0.287 *** (2.833)	0.253** (2.466)	0.247 ** (2.401)
ISU			0.377 *** (5.224)	0.366 *** (5.085)	0.357 *** (4.906)	0.354 *** (4.875)	0.329 *** (4.406)
UIL				0.241 *** (3.613)	0.236 *** (3.490)	0.229 *** (3.383)	0.235 *** (3.475)
Open					0.141 *** (2.866)	0.141 *** (2.853)	0.140 *** (2.847)
FP						0.092 ** (2.158)	0.105 ** (2.403)
EIL							0.081 * (1.813)
_cons	0.066 *** (200.741)	0.065 *** (77.140)	0.048 *** (14.365)	0.044 *** (12.511)	0.044 *** (12.450)	0.040 *** (10.276)	0.037 *** (9.047)
Year FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
N	3653	3653	3653	3653	3653	3653	3653
R ²	0.659	0.659	0.662	0.663	0.663	0.664	0.664

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. t-statistics are reported in parentheses.

4.2. Robustness Tests

4.2.1. Replacing the Explained Variable

Given that a single indicator can hardly capture the complex connotations of urban–rural integrated development, this study constructs an urban–rural integrated development indicator system across four dimensions: economic integration, social integration, ecological integration, and spatial integration [16] (see Table 4). The entropy-TOPSIS method is applied to compute a composite index (URI1), which replaces the original explained variable for re-estimation. As shown in Column (1) of Table 5, the coefficient on DF remains positive and statistically significant at the 1% level, confirming that the core findings are robust.

Table 4. Urban–rural integrated development indicator system.

Criterion Layer	Indicator Layer	Specific Measurement
Economic Integration	Economic development level	Gross Regional Product/Total population (Primary industry output/Primary industry employment)/(Secondary and tertiary industry output/Secondary and tertiary industry employment)
	Dual structure contrast coefficient	
	Urban–rural income contrast coefficient	Rural per capita disposable income/Urban per capita disposable income
	Urban–rural consumption contrast coefficient	Rural per capita consumption expenditure/Urban per capita consumption expenditure
Social	Urban–rural medical service supply	Number of beds in medical institutions/Total population

Integration	Urban–rural education investment	Public education expenditure/Fiscal expenditure
	Urban–rural medical condition gap	Medical institution beds per 10,000 people in county-level regions/Medical institution beds per 10,000 people in municipal districts
	Urban–rural education condition gap	Primary and secondary school students in county-level regions/Primary and secondary school students in municipal districts
Ecological Integration	Urban–rural air quality	Average PM2.5 mass concentration in urban and rural areas
	Urban–rural vegetation coverage	Mean vegetation index in urban and rural areas
	Urban–rural vegetation coverage contrast coefficient	Vegetation index in county-level regions/Vegetation index in municipal districts
	Urban–rural air quality contrast coefficient	Average PM2.5 concentration in county-level regions/Average PM2.5 concentration in municipal districts
Spatial Integration	Urban–rural transportation	Highway passenger volume/Total population
	Urban–rural communication connectivity	Number of mobile phone subscribers
	Permanent population urbanization rate	Urban permanent population/Total permanent population
	Urban–rural area contrast coefficient	Ratio of cultivated land area to municipal district area

Table 5. Results of Robustness Tests.

	(1)	(2)	(3)	(4)
	URI1	URI	URI	URI
DF	0.071 *** (3.512)	0.141 *** (3.904)		0.569 ** (2.283)
DF1			0.063 *** (3.920)	
Controls	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
City FE	YES	YES	YES	YES
N	3653	3653	3653	3653
R ²	0.663	0.639	0.639	0.653

Note: *** and ** denote statistical significance at the 1% and 5% levels, respectively. t-statistics are reported in parentheses.

4.2.2. Data Tail Reduction

To address the concern that winsorization at the 1% level may not adequately mitigate the influence of extreme values in the sample, this paper further applies two-tailed winsorization to all variables at the 5% level. The results are shown in Column (2) of Table 5. The estimated coefficient of the core explanatory variable, DF, is 1.141, which is positive and significant at the 1% level. This indicates that the promoting effect of data elements application on urban–rural integrated development remains robust after more stringently controlling for extreme values, further corroborating the reliability of the baseline regression results.

4.2.3. Replacing the Core Explanatory Variable

To reduce potential bias arising from indicator construction methodologies, this study replaces the proxy for the core explanatory variable with an alternative measure derived from the entropy-TOPSIS method and then re-estimates the model. As presented in Column (3) of Table 5, the coefficient on DF1 is positive and significant at the 1% level, consistent in direction with the baseline regression and exhibiting a higher level of statistical significance. This demonstrates that the promoting effect of data elements on urban–rural integration remains statistically significant across different measurement approaches, reinforcing the robustness of the conclusions.

4.2.4. Instrumental Variable Approach

To alleviate potential reverse causality and omitted variable bias associated with data elements application, given data availability, this study selects the annual number of newly added 2G base stations at the prefecture level as the instrumental variable and conducts two-stage least squares (2SLS) estimation. On the one hand, serving as a fundamental communication infrastructure, the coverage of 2G base stations directly dictates the convenience of regional data collection and transmission, exhibiting a high degree of correlation with data elements application and thereby satisfying the relevance requirement. On the other hand, concurrent with the iterative upgrading of communication technologies, the utility of 2G base stations has gradually diminished; their spatial distribution is predominantly determined by early policy orientations and historical investment layouts, bearing no direct nexus with contemporary urban–rural integrated development, which fulfills the exogeneity condition. Column (4) of Table 5 shows that under instrumental variable estimation, the estimated coefficient of DF is 0.569 and is positively significant at the 5% level. This aligns with the baseline regression findings, providing further causal evidence that data elements application fosters urban–rural integrated development.

4.3. Heterogeneity Analysis

4.3.1. Regional Heterogeneity Analysis

To investigate whether the impact of data elements application on urban–rural integrated development exhibits regional heterogeneity, the sample is partitioned into two distinct sets: coastal regions and inland regions, and municipalities directly under the central government, and other non-municipality cities. The estimation results are presented in Table 6. The results indicate that data elements application exerts a more significant effect in facilitating urban–rural integrated development in coastal regions and municipalities directly under the central government, both significant at the 1% level with coefficients of 0.189 and 0.543, respectively. Relative to inland regions, coastal regions possess distinct advantages in innovation capacity, economic growth, industrial structure optimization, and the degree of openness. These advantages provide favorable conditions for the deep integration of data elements with traditional production factors, thereby generating stronger diffusion effects through the extension of industrial, supply, and innovation chains. Concurrently, compared to other non-municipality cities, municipalities hold a first-mover advantage in infrastructure level, public service provision, policy resource allocation, and entrepreneurial support. This advantage cultivates a more conducive environment for the collection, processing, refining, annotation, categorization, and application of data elements, enabling them to play a more substantial role in driving urban–rural integrated development.

Table 6. Regional Heterogeneity Analysis Results.

	(1)	(2)	(3)	(4)
	Coastal Regions	Inland Regions	Municipalities	Other Cities
DF	0.189 *** (3.593)	0.047 (0.927)	0.543 *** (5.067)	0.051 (1.139)
Controls	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
City FE	YES	YES	YES	YES
N	1521	2132	52	3601
R ²	0.480	0.740	0.688	0.646

Note: *** denotes statistical significance at the 1% level. t-statistics are reported in parentheses.

4.3.2. Heterogeneity Analysis Based on Regional Characteristics

Disparities in regional economic growth, industrial structure advancement, and human capital levels engender pronounced heterogeneity in the role of data elements application in driving urban–rural integrated development. Economic growth is measured by gross regional product; industrial structure advancement is proxied by the ratio of the value added of the tertiary industry to that of the secondary industry; and human capital level is proxied by the ratio of the number of students enrolled in regular higher education institutions to the average annual population. As reported in Table 7, the facilitating effect of data elements application on urban–rural integrated development is more significant in regions characterized by superior economic growth, more advanced industrial structures, and abundant human capital. This is primarily attributable to the fact that regions with higher levels of economic growth furnish the material and market foundations requisite for the application of data elements; the advancement of industrial structures enhances the synergistic efficiency between data elements and traditional production factors; and abundant human capital bolsters the regional capacities to absorb, transform, and innovate using data elements. These factors collectively enable data elements application to fully actualize its potential value, thereby exerting a stronger impetus on urban–rural integrated development.

Table 7. Heterogeneity Analysis Based on Regional Characteristics Results.

	(1)	(2)	(3)	(4)	(5)	(6)
	Economic Growth		Industrial Structure Advancement		Human Capital Levels	
	Low	High	Low	High	Low	High
<i>DF</i>	−0.009 (−0.152)	0.180 *** (3.824)	0.048 (0.801)	0.159 *** (3.498)	0.043 (0.341)	0.139 *** (3.366)
Controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
<i>N</i>	1820	1833	1820	1833	1820	1833
<i>R</i> ²	0.728	0.590	0.652	0.675	0.690	0.600

Note: *** denotes statistical significance at the 1% level. *t*-statistics are reported in parentheses.

4.4. Spatial Spillover Effects

Given the near-zero marginal transmission cost and strong cross-regional mobility of data elements, their enabling effects are not confined by administrative boundaries, naturally exhibiting cross-regional spatial spillovers. Meanwhile, Chinese cities are inextricably linked geographically and economically, causing urban–rural integration to exhibit distinct spatial network characteristics. Therefore, this study employs LM and LR tests to select the Spatial Durbin Model for further analysis of the spatial spillover effects of data elements application on urban–rural integrated development [47]. The specific model is specified in Equation (3):

$$URI_{it} = \alpha_0 + \rho WURI_{it} + \alpha_1 DF_{it} + \varphi_1 WDF_{it} + \sum_j \alpha_j Control_{it} + \sum_j \varphi_j WControl_{it} + \eta_i + v_t + \varepsilon_{it} \quad (3)$$

In the above equation, *W* represents the spatial weight; ρ denotes the spatial autoregressive coefficient; while φ_1 and φ_j are the variables coefficients, respectively. Other variables remain consistent with those in Equation (1). A geographical adjacency matrix is constructed as the weight matrix [48].

Table 8 presents the estimation results for the spatial effect analysis. The spatial autoregressive coefficient *rho* is 0.517 and is statistically significant at the 1% level, indicating

that urban–rural integrated development exhibits a spatial clustering phenomenon. Among the main effects, the coefficient for data elements application is 0.832 and is statistically significant at the 1% level, indicating that data elements application has a significant positive impact on urban–rural integrated development. However, the coefficient for the spatial lag term is −0.671 and is statistically significant at the 1% level, suggesting that the development of data elements application in neighboring regions has a negative impact on local urban–rural integrated development. The underlying mechanism may stem from the siphon effect triggered by the application of data elements in neighboring regions. On the one hand, regions with higher levels of data elements application establish a first-mover advantage in digital industry development. According to industrial agglomeration theory, they can continuously attract data resources and human capital from surrounding areas, thereby exerting an adverse impact on the factor endowments and development momentum of those peripheries. On the other hand, regions with lower levels of data elements application exhibit notable deficiencies in data infrastructure and digital development, making it difficult to effectively absorb or transform potential data elements spillovers from adjacent areas. Furthermore, institutional factors across regions, such as local protectionism and administrative barriers, further impede the free cross-regional flow of data elements, thereby weakening or even offsetting potential positive spatial spillovers.

Table 8. Analysis of Spatial Spillover Effect Results.

	(1)	(2)	(3)	(4)
	Main	Wx	Spatial	Variance
DF	0.214 *** (2.609)	−0.926 * (−1.923)		
Controls	YES	YES		
rho			0.405 *** (3.381)	
sigma2_e				0.000 *** (42.761)
Year FE			YES	
City FE			YES	
Observations			3653	
R-squared			0.027	

Note: *** and * denote statistical significance at the 1% and 10% levels, respectively. Z-statistics are reported in parentheses.

5. Analysis of Transmission Channels

5.1. Mediating Model Construction

Characterized by non-scarcity and high permeability, data elements empower urban–rural integrated development not in isolation, but rather through a mutually embedded and symbiotic relationship with traditional production factors. To unpack the internal causal black box of how data elements drive urban–rural integrated development, this study introduces an analysis of the mechanism [49,50]. Equations (4) and (5) are constructed to test the mediating effects of data elements application on urban–rural integrated development through three transmission channels: facilitating labor mobility, mitigating capital misallocation, and enhancing TFP. The models are specified as follows:

$$\text{Med}_{it} = \beta_0 + \beta_1 \text{DF}_{it} + \sum_j \beta_j \text{Control}_{it} + \eta_i + \nu_t + \varepsilon_{it} \quad (4)$$

$$URI_{it} = \theta_0 + \theta_1 Med_{it} + \sum_j \theta_j Control_{it} + \eta_i + v_t + \varepsilon_{it} \quad (5)$$

In the above models, Med_{it} denotes the mechanism variables to be tested, specifically labor mobility, capital misallocation, and TFP.

5.2. Mediating Variables

5.2.1. Labor Mobility

Labor mobility is measured using the gravity model approach [51,52]. The underlying principle is that the shorter the geographic distance between economic agents, the greater the probability of interaction, and the continuous expansion of economic scale strengthens a region's attractiveness to factor inflows. This study introduces regional wage levels and housing price levels as gravitational variables for labor mobility and constructs an output-constrained gravity model to capture both the intensity and the direction of interregional labor flows, as specified in Equation (6):

$$LM_{ij} = \frac{L_i \cdot L_j \cdot \left(\frac{Wage_i}{Wage_j}\right) \cdot \left(\frac{House_i}{House_j}\right)}{D_{ij}^b} \quad (6)$$

In Equation (4), LM_{ij} denotes the quantity of labor flowing from region i to region j ; L_i and L_j represent the resident population of region i and region j , respectively; $Wage_i$ and $Wage_j$ denote the average wages of employed persons in urban units; $House_i$ and $House_j$ denote the average residential sales prices; D_{ij}^b is the geographic distance between region i and region j ; and b is the distance decay parameter, which is set to 2 in this study. Furthermore, the total labor mobility of region i , denoted as LM_i , is calculated as specified in Equation (7):

$$LM_i = \sum_{j=1}^m LF_{ij} \quad (7)$$

5.2.2. Capital Misallocation

Capital misallocation is measured based on the degree of capital price distortion [53]. Assuming that the production function varies across regions, the production function of a representative firm in region i is specified as $Y_i = A_i K_i^{\alpha_i} L_i^{\beta_i}$, which satisfies the condition of constant returns to scale. Here, Y_i denotes real output, K_i denotes the capital stock, L_i denotes the number of workers, α_i and β_i represent the capital and labor income shares, respectively. Under perfect competition, the product price is denoted by P_i , and the factor price of capital is R_i . If factor market distortions exist, the "distortionary taxes" faced by capital are denoted by τ_{k_i} , implying that firms actually pay $R_i(1 + \tau_{k_i})$ for capital. Based on the first-order condition for profit maximization of the representative firm and incorporating cross-regional resource constraints, the equilibrium quantity of capital allocated to region i is derived as shown in Equation (8):

$$K_i = \left[\frac{P_i \alpha_i Y_i}{(1 + \tau_{k_i}) R_i} / \sum_{j=1}^N \frac{P_j \alpha_j Y_j}{(1 + \tau_{k_j}) R_j} \right] K \quad (8)$$

The absolute factor distortion is defined as $\gamma_{k_i} = 1/(1 + \tau_{k_i})$, while the relative distortion captures the deviation of a region's factor price from the average distortion level across all regions. Accordingly, the relative capital distortion for region i is expressed as in Equation (9):

$$\theta_{k_i} = \gamma_{k_i} / \sum_{j=1}^N \left(\frac{S_j \alpha_j}{\alpha} \right) \gamma_{k_j} \quad (9)$$

where s_i denotes the share of the region i 's output in the total economy at equilibrium, and α is the output-weighted average capital contribution across regions. Solving the above yields the degree of capital misallocation, θ_{ki} , denoted as CD .

5.2.3. Total Factor Productivity

TFP is measured using the stochastic frontier analysis approach [54]. Real gross regional product is employed as the output indicator, while the number of employed persons and the capital stock serve as input indicators, with the capital stock estimated using the perpetual inventory method.

5.3. Test of Transmission Channels

To further verify the transmission mechanisms through which data elements application drives urban–rural integrated development—namely, facilitating labor mobility, mitigating capital misallocation, and enhancing TFP—this study conducts mechanism tests, with the results reported in Table 9.

Columns (1)–(2) report the test results for the labor mobility channel. Column (1) estimates the impact of data elements application on labor mobility; the coefficient is 0.013 and statistically significant at the 1% level. This indicates that, leveraging its high permeability, the data elements application effectively reduces information asymmetry in urban and rural labor markets and broadens the job search radius, thereby promoting the cross-regional flow of labor. Column (2) further verifies the promoting effect of labor mobility on urban–rural integrated development, with a coefficient of 0.653 that is statistically significant at the 1% level. This confirms that data elements application advances urban–rural integrated development by facilitating labor mobility, which validates Hypothesis 2a.

Columns (3)–(4) present the results for the capital misallocation channel. Column (3) shows the impact of data elements application on capital misallocation; the coefficient is negative and statistically significant at the 1% level, indicating that data elements enhance the information transparency and matching efficiency of financial markets, guiding capital to flow toward highly efficient sectors and thereby reducing the level of capital misallocation. Combined with the results in Column (4), the improvement in capital misallocation makes a significant positive contribution to urban–rural integrated development. Therefore, Hypothesis 2b is supported.

Columns (5)–(6) report the mediation results of TFP. In Column (5), the coefficient of data elements application on TFP is 0.309 and statistically significant at the 1% level. This confirms that data elements, acting as a general-purpose technology, generate technological spillovers and scale effects, effectively enhancing the TFP of the economic system. Combined with Column (6), TFP improvement plays a statistically significant mediating role in the nexus between data elements application and urban–rural integrated development, which validates Hypothesis 2c.

Table 9. Test Results of Transmission Channels.

	(1)	(2)	(3)	(4)	(5)	(6)
	LM	URI	CD	URI	TFP	URI
DF	0.013 *** (2.791)		−0.034 *** (−4.769)		0.309 *** (4.838)	
LM		0.653 *** (4.091)				
CD				−0.245 *** (−3.039)		
TFP						0.040 ***

						(3.906)
Controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
N	3653	3653	3653	3653	3653	3653

Note: *** denotes statistical significance at the 1% level. t-statistics are reported in parentheses.

6. Discussion and Conclusions

6.1. Discussion

The estimation results from the two-way fixed-effects model indicate that the application of data elements significantly promotes urban–rural integrated development. This finding is broadly consistent with the core conclusions of the existing literature. Studies on the digital economy have established that the widespread penetration of digital technologies can reduce information asymmetry and enhance factor allocative efficiency, thereby facilitating urban–rural integrated development [55]. However, most existing research adopts a macro perspective centered on the digital economy or digitalization, and has yet to thoroughly examine the role of data factors as an independent and novel production factor in shaping urban–rural integrated development. This study disentangles data elements from the macro context of the digital economy, reveals their unique advantages from three dimensions—cross-boundary circulation, multidimensional integration, and efficiency gains from reuse—and empirically tests their promotive effect on urban–rural integrated development.

We construct and examine three transmission channels through which data elements application affects urban–rural integrated development: facilitating labor mobility, mitigating capital misallocation, and enhancing TFP. Regarding the labor mobility channel, existing studies predominantly adopt a unidirectional perspective on urban–rural labor migration, focusing on the transfer of rural surplus labor to urban areas [56]. This study further reveals that data elements not only lower information barriers in urban–rural labor markets but also expand the allocative space for labor through emerging employment forms such as the platform economy and flexible employment, enabling smoother and more efficient two-way labor mobility between urban and rural areas. With respect to the capital misallocation channel, extant literature centers on the accessibility of urban–rural financial services and the equilibrium of credit rationing [57]. By contrast, this study highlights how the application of data elements mitigates urban–rural capital misallocation by enhancing information transparency and reducing transaction costs. As for the TFP channel, prior research typically treats TFP as an outcome variable of digital economy development rather than a transmission mechanism [58]. This study incorporates TFP into a mediation effect framework, empirically verifying the transmission path whereby data elements application enhances TFP via efficiency gains from reuse, which in turn promotes urban–rural integrated development.

Regarding heterogeneity, most extant literature delineates heterogeneity along the macro geographic gradient of eastern, central, and western China [59]. This study further incorporates intra-city structural characteristics and development endowments into the explanatory framework. The results show that the promoting effect of data elements application on urban–rural integrated development is more significant in regions characterized by higher levels of economic development, more advanced industrial structures, and more abundant human capital. This finding corroborates the established argument concerning the digital divide in digital economy development [60]. It indicates that the empowering effect of data elements can only reach its optimal level through deep integration between data elements and traditional factors when a region's factor endowments have

accumulated to a certain threshold. This provides empirical support for implementing differentiated digital village strategies tailored to local conditions and specific contexts.

In terms of spatial spillover effects, this study finds that while data elements application facilitates local urban–rural integrated development, it imposes a significant negative impact on urban–rural integration in neighboring regions. This finding aligns with the theoretical expectations of polarization and peripheral effects in New Economic Geography: in the initial stage of market-oriented data elements allocation, regions with first-mover advantages may generate adverse impacts on neighboring areas through factor attraction. Most existing studies on factor allocation rely on traditional non-spatial econometric models, which tend to neglect the spatial correlation of cross-regional factor flows [61]; others unilaterally emphasize the inherently positive spatial spillovers of the digital economy [62]. This study demonstrates that against the backdrop of regional integration and marketization, the high mobility of data elements may accelerate the agglomeration of high-quality labor and capital toward core cities or advantaged regions, thereby widening the urban–rural gap relative to surrounding less-developed areas. Therefore, while harnessing the advantages of data elements, it is imperative to remain vigilant about the inter-regional development imbalances they may induce. Establishing cross-regional collaborative governance and benefit-sharing mechanisms is of paramount importance.

6.2. Conclusions

Drawing on prefecture-level panel data from 281 cities in China from 2011 to 2023, this study employs a two-way fixed-effects model to examine the effects and transmission channels of data elements application on urban–rural integrated development. Four main conclusions are drawn from the analysis. (1) Data elements application exerts a significant and positive promotional effect on urban–rural integrated development. On average, a 1% increase in the level of data elements application is associated with a 0.089% rise in the urban–rural integrated development index. This effect remains statistically significant following the stepwise inclusion of control variables and a series of robustness checks, further corroborating the positive role of data elements application in advancing urban–rural integrated development. (2) The mechanism analysis demonstrates that data elements application advances urban–rural integrated development through three primary transmission channels: promoting labor mobility, mitigating capital misallocation, and elevating TFP. (3) Heterogeneity analysis reveals that the promotive effect of data elements application on urban–rural integrated development exhibits both regional heterogeneity and heterogeneity in local development endowments. In terms of regional heterogeneity, compared with inland regions and other non-municipality cities, data elements application plays a more significant role in promoting urban–rural integrated development in coastal regions and municipalities directly under the central government. Regarding local characteristic heterogeneity, the promotional effect is more pronounced in regions characterized by higher levels of economic development, more advanced industrial structures, and more abundant human capital. (4) Analysis of spatial spillover effects indicates that the application of data elements exhibits significant agglomeration characteristics across space. However, the coefficient of data elements application in the spatial lag term stands at -0.926 and is statistically significant at the 10% level, implying that data elements application in neighboring regions may exert a negative impact on local urban–rural integrated development via the siphon effect.

This study makes three key contributions to the extant literature. First, while most existing studies treat data elements as an implicit component of the digital economy, this study explicitly defines the unique attributes of data elements and empirically examines their direct effect on urban–rural integrated development, thereby clarifying the specific role of data elements in shaping urban–rural integration. Second, by identifying three

distinct transmission channels through which data elements affect urban–rural integrated development, this study extends traditional factor allocation theory. Rather than simply substituting for traditional input factors, data elements function as a catalyst that improves the allocation efficiency of existing resources between urban and rural areas, highlighting that data elements play a multiplicative rather than additive role in advancing urban–rural integration. Third, the finding of significant negative spatial spillovers constitutes an important complement to the spatial economic theory of urban–rural relations. This study shows that data elements application in one region may adversely affect urban–rural integrated development in neighboring regions through the siphon effect, revealing the dual nature of data-driven development as both centripetal and centrifugal forces within a regional system. This provides a more comprehensive understanding of the spatial dimension of data-driven development.

6.3. Policy Recommendations

6.3.1. Accelerate the Construction of Data Elements Markets and Establish an Urban–Rural Integrated Data Element Circulation System

For the central government, the top priority is to advance the construction of a unified and open national data elements market. Efforts should focus on exploring and refining foundational institutional frameworks concerning the confirmation of data property rights, circulation and trading, revenue distribution, and security governance. This ensures that throughout the entire process of data collection, circulation, and utilization, rights and responsibilities are clearly defined, transactions are efficient, and security is maintained. For local governments, the emphasis should be placed on establishing regional data trading platforms, with particular priority given to promoting the standardization, assetization, and market-oriented circulation of agriculture-related data—such as meteorological conditions, soil data, agricultural product prices, and logistics information. This aims to break down data silos, facilitate the efficient flow of data elements between urban and rural areas, and achieve deep urban–rural integration at the industrial, resource, and informational levels.

6.3.2. Deepen the Integration of Data Elements with Labor, Capital, and Technology to Unleash the Multiplier Effect of Data Elements

First, construct a unified urban–rural digital platform covering employment information and vocational services. Integrating urban and rural labor force data with vocational skill information can improve information transparency and labor matching efficiency, thereby promoting the orderly flow of labor between urban and rural areas. Second, improve the digital inclusive financial system and further deepen the reform of the rural property rights system. Promoting institutional innovation in factor circulation and policy-based financial support can guide urban capital to flow into rural areas, mitigate urban–rural capital misallocation, and improve capital utilization efficiency. Third, accelerate the extensive application of digital technologies in agriculture, manufacturing, and service industries. Making full use of big data, artificial intelligence, the Internet of Things, cloud computing, and other digital technologies to optimize production processes, improve management efficiency, and facilitate technology diffusion will extend and integrate industrial chains, providing sustained impetus for urban–rural integrated development.

6.3.3. Implement Differentiated and Targeted Strategies to Promote Urban–Rural Integrated Development in a Localized Manner

Coastal regions and municipalities directly under the central government should give full play to their comparative advantages in digital economy development, industrial agglomeration, and technological innovation to explore higher-level models of urban–rural integrated development, realizing in-depth linkage of resources, technology, and industries between urban and rural areas. Inland regions and other non-municipality cities should increase fiscal transfer payments and policy inclination, systematically shore up weaknesses in digital infrastructure, improve human capital quality, and optimize industrial structures, so as to strengthen the empowering role of data elements application in facilitating the efficient flow of urban–rural factors. For regions with a higher level of economic development, advanced industrial structures, and abundant human capital, greater emphasis should be placed on amplifying the multiplier and demonstration effects of data elements. Data elements should be used to empower the extension of urban–rural industrial chains through facilitating interactive development between advanced urban industries and rural characteristic industries, and achieve synergies between industrial structure upgrading and coordinated regional economic development. Meanwhile, optimize the flow mechanism of high-quality human capital between urban and rural areas, encourage innovative factors to extend to grassroots regions, and promote the two-way diffusion of knowledge, technology, and innovative resources, ultimately achieving high-quality urban–rural integrated development.

6.3.4. Mitigate the Spatial Siphon Effect and Promote Coordinated Regional Development

First, a cross-regional data sharing mechanism should be established to facilitate the cross-regional circulation of data elements. The government should design an institutional framework mandating data sharing among adjacent regions and create an incentive structure to reward areas that contribute to collective data resources. Cities should establish regional data alliances and collaborative governance platforms to promote the equitable distribution of data-driven development benefits across jurisdictions. Second, digital infrastructure construction in less developed areas should be strengthened to narrow the digital divide and enhance their absorptive capacity for positive spillovers from neighboring regions. Targeted investments should be directed toward broadband networks, cloud computing facilities, and data processing centers in underdeveloped areas, ensuring they possess the technological foundation to benefit from data-driven innovations in developed regions. Third, a mechanism for the coordinated development of the regional digital economy should be established to prevent vicious competition among regions and foster mutually beneficial outcomes. This includes creating cross-regional digital economy coordination committees, formulating joint digital industry planning frameworks, and implementing complementary digital infrastructure investment strategies.

6.4. Limitations and Future Research

While this study provides robust empirical evidence on the relationship between data elements application and urban–rural integrated development, several limitations remain that warrant further exploration in future research. Regarding data availability and measurement, the construction of the data elements application index relies on publicly available prefecture-level statistical data, which may not fully capture the depth and quality of data utilization within cities. Future research should endeavor to develop more refined measurement frameworks that incorporate firm-level and household-level indicators of data elements utilization, as well as qualitative indicators capturing data governance quality and institutional dimensions. Regarding the comprehensiveness of the mechanism analysis, although the mediating effect model provides valuable evidence for the

three transmission channels identified in this study, it does not exhaust the full spectrum of transmission mechanisms. Additional pathways—such as data-driven equalization of public services, improvements in digital governance, and the expansion of rural consumer markets—may also play pivotal roles. Future research should examine these additional mechanisms and employ structural equation modeling or integrated analytical frameworks to capture the full complexity of the causal chains through which data elements shape urban–rural integration. Furthermore, as the market for data elements continues to evolve and mature in China and other regions, longitudinal studies tracking the dynamic effects of data elements application on urban–rural integrated development will contribute to a more comprehensive understanding of the long-term impacts of data-driven development on spatial inequality and regional convergence.

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

Funding: This research was funded by the National Social Science Fund of China, grant number 25XKS015; Shaanxi Provincial Social Science Fund Project, grant number 2025A001.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Niu, B.; Ge, D.Z.; Sun, J.W.; Sun, D.Q.; Ma, Y.Y.; Ni, Y.L.; Lu, Y.Q. Multi-scales urban-rural integrated development and land-use transition: The story of China. *Habitat Int.* **2023**, *132*, 102744.
2. Liu, Y.S.; Zang, Y.Z.; Yang, Y.Y. China's rural revitalization and development: Theory, technology and management. *J. Geogr. Sci.* **2020**, *30*, 1923–1942.
3. Wang, S.B.; Song, J.Y. Fintech and urban-rural integrated development: A human capital perspective. *Financ. Res. Lett.* **2026**, *92*, 109534.
4. Liao, R.F.; Wei, Y.H.; Bai, Y.; Liu, J.X. Bridging the divide: How agricultural technological innovation narrows the urban-rural income gap in China. *Front. Sustain. Food Syst.* **2025**, *9*, 1595161.
5. Tong, L.; Tang, L. Regional Differences and Spatial-Temporal Evolution Characteristics of Digital Economy Development in China. *Sustainability* **2025**, *17*, 4737.
6. Sun, G.L.; Ning, C.T.; Chen, C.; Li, M.D. Data elements and enterprise total factor productivity: The role of innovation and resource allocation. *Int. Entrep. Manag. J.* **2026**, *22*, 102.
7. Gartlehner, G.; Kahwati, L.; Hilscher, R.; Thomas, I.; Kugley, S.; Crotty, K.; Viswanathan, M.; Nussbaumer-Streit, B.; Booth, G.; Erskine, N.; et al. Data extraction for evidence synthesis using a large language model: A proof-of-concept study. *Res. Synth. Methods* **2024**, *15*, 576–589.
8. Mikalef, P.; Boura, M.; Lekakos, G.; Krogstie, J. Big data analytics and firm performance: Findings from a mixed-method approach. *J. Bus. Res.* **2019**, *98*, 261–276.
9. Yang, L.; Zuo, Y.X. Can the Input of Data Elements Improve Manufacturing Productivity? Effect Measurement and Path Analysis. *IEEE Trans. Eng. Manag.* **2024**, *71*, 15320–15332.
10. Jiang, J.L. The multiplier effect of data elements empowering green economic development: A mechanism examination based on innovation-driven. *Appl. Econ.* **2025**, *57*, 11122–11137.
11. Jing, R.F.; Xu, Y.M.; Liu, R.Z. Data element, big data and enterprise innovation: Evidence from data trading platforms and national big data comprehensive pilot zones. *Int. Rev. Financ. Anal.* **2025**, *105*, 104426.

12. Li, B.Q.; Zheng, J.; Jiang, J. The Impact of Data Element Marketization on Labor Income Share: Quasi-Natural Experiment Evidence from the Establishment of Data Trading Platforms in China. *Emerg. Mark. Financ. Trade* **2025**, *62*, 3040–3057.
13. Zhang, Z.; Yu, Y. The nonlinear impact of data factor flow on coordinated regional economic development: Based on the moderating effect of digital infrastructure. *Financ. Res. Lett.* **2025**, *85*, 108044.
14. Yan, D.S.; Sun, W.; Li, P.X.; Liu, C.G.; Li, Y.J. Effects of economic growth target on the urban-rural income gap in China: An empirical study based on the urban bias theory. *Cities* **2025**, *156*, 105518.
15. Zhu, B.; Zhang, S.; Chen, W.; Shi, H.; Wang, Y. The positive impact of urban-rural integration on consumption gap among residents in China. *Cities* **2026**, *168*, 106454.
16. Li, L.; Yang, L.; Wang, H.; Yang, S.; Yan, Z.H. Spatiotemporal Evolution Characteristics and Impact Factors of Urban-rural Integrated Development in China. *Chin. Geogr. Sci.* **2025**, *35*, 802–818.
17. Zhai, G.; Zhang, M.X.; Hu, Y.H.; Chen, Y.W.; Wu, C.F.; Lu, Y.P.; He, T.T. Nighttime light metrics for analysing urban-rural economic disparities: A case study in 36 Chinese metropolitan areas. *Cities* **2025**, *162*, 105930.
18. Tan, X.J.; Kamaruddin, R.B.; Hu, S.Y.; Peng, L.S.; Que, Y.X.; Cai, W.W. Rural revitalization and urban-rural income gap: A perspective from land transfer scale. *Financ. Res. Lett.* **2025**, *83*, 107705.
19. Li, K.F.; Tian, W.F.; Wang, S.K.; Li, Z.H. Does Digital Inclusive Finance Promote Urban-Rural Integration? *Emerg. Mark. Financ. Trade* **2026**, *62*, 1411–1432.
20. Yang, M.; Sun, H.L. The impact of digital inclusive finance on urban-rural income gap. *Financ. Res. Lett.* **2026**, *92*, 109572.
21. Bi, J.Y. Can rural areas in China be revitalized by digitization? A dual perspective on digital infrastructure and digital finance. *Financ. Res. Lett.* **2024**, *67*, 105753.
22. Liang, K.L.; Liu, W.Q. Digital infrastructure and urban-rural income gap: Empirical evidence from China. *Int. Rev. Econ. Financ.* **2026**, *107*, 105003.
23. Lei, J.C.; Xie, Y.; Chen, Y.S.; Zhong, T.Y.; Lin, Y.C.; Wang, M. The Transformation of Peri-Urban Agriculture and Its Implications for Urban-Rural Integration Under the Influence of Digital Technology. *Land* **2025**, *14*, 375.
24. Arrow, K.J.; Debreu, G. Existence of an Equilibrium for a Competitive Economy. *Econometrica* **1954**, *22*, 265–290.
25. Gao, D.Y.; Liu, C.; Sun, Z.W. Data elements and corporate innovation: A discussion of corporate innovation strategy. *Financ. Res. Lett.* **2025**, *76*, 106970.
26. Wang, C.C.; Chiou, J.R. Optimal privatization policy in the presence of network externalities. *Oxf. Econ. Pap.* **2015**, *67*, 514–527.
27. Schmitz, P.W. Information Gathering, Transaction Costs, and the Property Rights Approach. *Am. Econ. Rev.* **2006**, *96*, 422–434.
28. Rao, N. Asymmetric information and search frictions: A neutrality result. *Econ. Lett.* **2016**, *147*, 138–141.
29. Dong, L.; Zhu, X.W.; Yang, L.W.; Jiang, G.L. Unleashing the power of data element markets: Driving urban green growth through marketization, innovation, and digital finance. *Int. Rev. Econ. Financ.* **2025**, *99*, 104070.
30. Selod, H.; Shilpi, F. Rural-urban migration in developing countries: Lessons from the literature. *Reg. Sci. Urban Econ.* **2021**, *91*, 103713.
31. Li, G.Q.; Qin, J.H. Income effect of rural E-commerce: Empirical evidence from Taobao villages in China. *J. Rural Stud.* **2022**, *96*, 129–140.
32. Lu, H.Y.; Zhao, P.J.; Hu, H.Y.; Zeng, L.G.; Wu, K.S.; Lv, D. Transport infrastructure and urban-rural income disparity: A municipal-level analysis in China. *J. Transp. Geogr.* **2022**, *99*, 103292.
33. Adamopoulos, T.; Brandt, L.; Leight, J.; Restuccia, D. Misallocation, Selection, and Productivity: A Quantitative Analysis With Panel Data From China. *Econometrica* **2022**, *90*, 1261–1282.
34. Lu, Y.C.; Zhuang, J.K.; Yang, C.L.; Li, L.; Kong, M. How the digital economy promotes urban-rural integration through optimizing factor allocation: Theoretical mechanisms and evidence from China. *Front. Sustain. Food Syst.* **2025**, *9*, 1494247.
35. Comin, D.; Mestieri, M. If Technology Has Arrived Everywhere, Why Has Income Diverged? *Am. Econ. J. Macroecon.* **2018**, *10*, 137–178.
36. Akerberg, D.A.; Caves, K.; Frazer, G. Identification Properties Of Recent Production Function Estimators. *Econometrica* **2015**, *83*, 2411–2451.
37. Ren, Y.H.; Zhang, J.; Wang, X. How does data factor utilization stimulate corporate total factor productivity: A discussion of the productivity paradox. *Int. Rev. Econ. Financ.* **2024**, *96*, 103681.
38. Jiang, S.; Zhou, J.; Qiu, S. Digital Agriculture and Urbanization: Mechanism and Empirical Research. *Technol. Forecast. Soc. Change* **2022**, *180*, 121724.
39. Pang, H.; Huang, Z.; Son, S.; Sha, X. Land-Finance Intensity, the Urban–Rural Income Gap, and Female Educational Attainment: A Mediation Analysis Based on Provincial Panel Data from China. *Land* **2026**, *15*, 673.

40. Shan, B.Y.; Zhang, Q.; Ren, Q.X.; Yu, X.W.; Chen, Y.Q. Spatial heterogeneity of urban-rural integration and its influencing factors in Shandong province of China. *Sci. Rep.* **2022**, *12*, 14317.
41. Janssen, M.; Van Der Voort, H.; Wahyudi, A. Factors influencing big data decision-making quality. *J. Bus. Res.* **2017**, *70*, 338–345.
42. Pan, W.; Wang, J.; Li, Y.R.; Chen, S.T.; Lu, Z. Spatial pattern of urban-rural integration in China and the impact of geography. *Geogr. Sustain.* **2023**, *4*, 404–413.
43. Han, J.; Jiang, S.C. Does the Adjustment of Industrial Structure Restrain the Income Gap between Urban and Rural Areas. *Economics* **2024**, *18*, 20220112.
44. Chen, B.; Zhu, H.S.; Dai, J.X. The impact of urban digital economy on urban innovation: The mediating role of trade openness. *Asian Geogr.* **2025**, *42*, 195–218.
45. Zhang, X.; Zhu, Y.; Hou, A. Fiscal decentralization and urban-rural inequality of income acquisition opportunities: Micro evidence from China. *Financ. Res. Lett.* **2025**, *85*, 107859.
46. Cao, S.; Yu, N.N.; Wu, Y.; Wang, Z.H.; Mi, J.N. The Educational Level of Rural Labor, Population Urbanization, and Sustainable Economic Growth in China. *Sustainability* **2020**, *12*, 4860.
47. Elhorst, J.P. Matlab Software for Spatial Panels. *Int. Reg. Sci. Rev.* **2014**, *37*, 389–405.
48. Peng, W.B.; Yin, Y.; Kuang, C.G.; Wen, Z.Z.; Kuang, J.S. Spatial spillover effect of green innovation on economic development quality in China: Evidence from a panel data of 270 prefecture-level and above cities. *Sustain. Cities Soc.* **2021**, *69*, 102863.
49. Sun, Y.M. Digital transformation and corporates' green technology innovation performance-The mediating role of knowledge sharing. *Financ. Res. Lett.* **2024**, *62*, 105105.
50. Liang, L.; Li, Y. How does government support promote digital economy development in China? The mediating role of regional innovation ecosystem resilience. *Technol. Forecast. Soc. Change* **2023**, *188*, 122328.
51. Haas, A.; Osland, L. Commuting, Migration, Housing and Labour Markets: Complex Interactions. *Urban Stud.* **2014**, *51*, 463–476.
52. Feng, Q.Y.; Chen, Z.H.; Cheng, C.C.; Chang, H.Q. Impact of high-speed rail on high-skilled labor mobility in China. *Transp. Policy* **2023**, *133*, 64–74.
53. Hao, Y.; Gai, Z.Q.; Wu, H.T. How do resource misallocation and government corruption affect green total factor energy efficiency? Evidence from China. *Energy Policy* **2020**, *143*, 111562.
54. Guo, B.S.; Yu, H.; Jin, G. Urban green total factor productivity in China: A generalized Luenberger productivity indicator and its parametric decomposition. *Sustain. Cities Soc.* **2024**, *106*, 105365.
55. Liu, F.; Wang, D.; Liu, C.; Tian, G. Data-driven coordination and optimization of the digital economy and urban-rural integration. *Energy Econ.* **2025**, *152*, 108959.
56. Gu, X.Y.; Liu, X.T. Uneven spatial patterns and disparate socioeconomic impacts of intercity labor mobility in China. *Ann. Reg. Sci.* **2025**, *74*, 57.
57. Zhang, W.W.; Xia, X.Y.; Zhu, Y.; Zhao, S.J.; Zhou, Z.J. Rural financial institutions, access to credit and production inputs of rural households. *Financ. Res. Lett.* **2024**, *70*, 106298.
58. Tu, J.; Wei, X.; Bin Razik, M.A. The impact of digital technology on total factor productivity in manufacturing enterprises. *Sci. Rep.* **2025**, *15*, 23543.
59. Lin, X.; Jiang, Y. The impact of industrial digitalization on the urban-rural income gap. *PLoS ONE* **2025**, *20*, e0335065.
60. Xu, Y.; Tao, J. Digital dividend or digital divide?—Evidence from China. *Appl. Econ.* **2025**, *57*, 5033–5048.
61. Chen, H.; Qin, S.; Wang, H. Spatial Spillover Effect and Nonlinear Relationship between Factor Misallocation and High-Quality Agricultural Economy: Evidence from 154 Cities in Major Grain-Producing Areas of China. *Agriculture* **2023**, *13*, 1922.
62. Liao, Y.; Tian, J.; Chang, X.; Wu, G.; Wang, B. Factor Mobility and Urban-Rural Integration in China: Unpacking Direct, Indirect, and Spatial Spillover Effects at the County Level. *Land* **2026**, *15*, 975.

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.