



Article

Observations on Inclusive Development of Rural E-Commerce in China: Traffic Resource Allocation Based on the DEMATEL-ISM Model

Yang He ¹, Qingjun Li ² and Jianxin Zhu ^{1,3,*}

¹ School of Economics and Management, Harbin Engineering University, Harbin 150001, China

² Heilongjiang Provincial Science and Technology Commission, Harbin 150040, China

³ Key Laboratory of Big Data and Business Intelligence Technology, Ministry of Industry and Information Technology, Harbin Engineering University, Harbin 150001, China

* Correspondence: heuzjx@hrbeu.edu.cn

Abstract

Amid the integration of the digital and platform economies, rural e-commerce traffic resource integration involves complex multi-factor interactions and hidden causal pathways, making key mechanism identification challenging. To overcome the limitations of conventional DEMATEL–ISM studies, which often rely on manually defined factors and lack integration between factor discovery and structural analysis, this study proposes an integrated LLM–DEMATEL–ISM workflow. Rather than a simple application of LLMs, the proposed framework combines LLM-assisted factor identification, structured validation, expert evaluation, and DEMATEL–ISM modeling into a unified analytical process. LLMs support multi-source factor discovery, while expert validation ensures conceptual reliability; DEMATEL and ISM further quantify causal relationships and reveal hierarchical transmission structures. The results identify market entity scale, organizational coordination capability, and digital traffic mechanisms as key hub factors. The main methodological contribution lies in establishing a reproducible human–machine–structure collaborative framework that bridges data-driven factor discovery and expert-validated structural modeling, providing a new analytical paradigm for complex-systems research in the digital economy.

Keywords: rural e-commerce; traffic resource integration; large language models; DEMATEL-ISM; causal structure analysis



Academic Editors: Carman Ka
Man Lee and Ting Chi

Received: 13 May 2026

Revised: 20 August 2026

Accepted: 23 August 2026

Published: 8 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

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

1. Introduction

Driven by national initiatives such as “Digital China” and the “Rural Revitalization Strategy,” rural e-commerce has emerged as a primary engine for economic transformation and agricultural modernization. By leveraging the disintermediation of digital platforms, rural producers can bypass traditional intermediaries and reach national markets at substantially lower transaction costs, fostering inclusive growth and strengthening household resilience [1,2].

As the platform economy matures, the competitive landscape of rural e-commerce has shifted fundamentally: market success no longer depends solely on physical distribution but is increasingly dictated by platform algorithms and digital traffic-allocation mechanisms [3]. In this algorithmic ecosystem, “traffic resources” have surpassed traditional factors of production to become the decisive strategic asset governing product visibility,

user engagement, and market competitiveness [4,5]. For rural participants—who often lack brand recognition and digital marketing capabilities—the efficiency and equity of traffic allocation have become central to their sustainable development. Analyzing how traffic resources are allocated is therefore essential to understanding the rural digital economy.

Prior scholarship has documented the transformative role of rural e-commerce, focusing largely on its socio-economic impacts and structural foundations—how digital infrastructure and supply-chain optimization contribute to household resilience and regional income growth [4,6–10]. Yet this literature centers on the expansion phase of rural e-commerce and pays limited attention to the algorithmic stage, in which “traffic resources” have become the decisive factor of production.

Within this context, the attention economy offers an essential lens for examining the micro-level mechanisms of resource allocation: in information-rich ecosystems, user attention is the scarce resource that determines economic value. A critical question therefore remains under-explored: how do platform-driven traffic-allocation mechanisms shape the allocative efficiency and economic equity experienced by different rural e-commerce participants? Does the prevailing logic of traffic distribution widen the digital divide, or does it open a more inclusive path for rural revitalization? Answering these questions is vital to understanding how digital platforms can sustainably empower rural economies.

To address these issues, this study develops an analytical framework that integrates Large Language Model (LLMs) semantic analysis with the DEMATEL-ISM method. LLMs first analyze academic literature, policy documents, and industry reports to identify the key factors governing traffic integration; the Decision-Making Trial and Evaluation Laboratory (DEMATEL) method then traces the causal relationships among these factors; Interpretive Structural Modeling (ISM) finally reveals the hierarchical mechanisms that shape traffic resource allocation.

This study makes three main contributions. First, it introduces the attention-economy perspective to reframe traffic as a scarce, strategic digital resource, thereby enriching the theoretical understanding of the rural digital economy. Second, it proposes an integrated LLM–DEMATEL–ISM workflow that extends conventional DEMATEL–ISM from a knowledge-driven to a data-driven hybrid framework. LLM-based semantic mining identifies candidate factors from heterogeneous textual sources—academic literature, policy documents, and industry reports—before DEMATEL–ISM modeling, reducing the subjective bias and semantic redundancy of manual screening and improving the comprehensiveness, transparency, and reproducibility of factor identification and structural analysis. Third, by uncovering the latent causal relationships and hierarchical transmission pathways among traffic-allocation factors, the study provides actionable policy insights for advancing inclusive digital development and more equitable resource allocation within China’s rural e-commerce.

1.1. Literature Review

Existing research on the integration of traffic resources in rural e-commerce primarily encompasses several research directions: rural e-commerce development and organizational capabilities, platform traffic distribution mechanisms, value co-creation in digital platforms, and complex-systems modeling approaches. As platform economies and artificial intelligence have advanced, scholars have begun re-examining e-commerce from the perspective of digital platform ecosystems, particularly how platform algorithms, attention-resource allocation, and data-driven decision-making reshape these ecosystems [11]. A multidimensional review of this literature is essential to building a theoretical framework for rural e-commerce traffic resource integration.

1.1.1. Research on Rural E-Commerce Development and Organizational Capabilities

Rural e-commerce represents an important manifestation of the integrated development of the digital economy and rural industries. With the proliferation of internet technology and the implementation of digital village strategies, e-commerce has gradually become a significant driving force for rural economic transformation. Numerous studies have analyzed, from a macro perspective, how rural e-commerce affects agricultural industrial restructuring, the circulation efficiency of agricultural products, and farmers' income. For instance, Liu et al. (2024) found, using data on China's rural e-commerce development, that digital platforms substantially reduce transaction costs and expand the market reach of agricultural products [9]. Research by Dong et al. (2024) indicates that rural e-commerce can increase agricultural product sales prices and boost farmers' income by reducing intermediate circulation links, thereby promoting inclusive growth in rural economies [7]. Regarding developmental mechanisms, scholars generally recognize infrastructure as a key driver of rural e-commerce. Studies indicate that the development of rural e-commerce is highly dependent on infrastructure conditions such as logistics networks, information and communication technology, and digital payment systems. Cold-chain logistics and warehousing-distribution networks are especially important for the online sale of fresh agricultural products. Furthermore, some studies analyze rural e-commerce development from an institutional environment perspective. For example, government policy support, e-commerce demonstration projects, and the construction of rural e-commerce service stations all play important roles in promoting the formation of rural e-commerce ecosystems.

As research has deepened, scholars have shifted focus from infrastructure to the capabilities of rural e-commerce actors and to industrial organization. These studies show that rural e-commerce development depends not only on external conditions but also on farmers' digital literacy, platform-operation capabilities, and organizational models [12]. For example, some studies point out that farmers' digital skills and e-commerce operational capabilities directly affect the online sales performance of agricultural products [13]. Additionally, organizational entities such as rural cooperatives, e-commerce service enterprises, and rural e-commerce leaders also play important roles in promoting farmers' participation in e-commerce activities. However, although existing research has achieved relatively rich results in the developmental mechanisms of rural e-commerce, most studies primarily focus on farmers' income growth or e-commerce sales performance, with insufficient attention paid to platform traffic acquisition and allocation mechanisms. However, although this literature is rich, most studies focus on farmers' income or sales performance and treat traffic as a passive outcome of sales rather than a strategic, integrable resource [14,15]. How rural entities proactively acquire and integrate traffic resources within algorithmic environments remains largely unexamined.

1.1.2. Research on Platform Traffic Distribution Mechanisms and Digital Ecosystems

As the digital platform economy has grown, traffic has become a core competitive resource in digital markets. Within platform ecosystems, platform enterprises configure user attention through search rankings, recommendation algorithms, and content distribution mechanisms, thereby affecting merchants' market exposure and sales conversion rates. Scholars have analyzed traffic resource allocation through attention-economy theory, which holds that in information-saturated environments human attention becomes the scarcest resource and platform algorithms govern its allocation [6,16].

E-commerce platforms use recommendation algorithms to analyze user behavior, product attributes, and content quality, delivering personalized recommendations and allocating traffic accordingly [17]. This data-driven recommendation mechanism, while enhancing user experience and platform efficiency, has also transformed traditional market

competition models. Research shows that on algorithm-driven platforms traffic distribution is often top-heavy: a few leading merchants capture most of the platform's traffic, while small and medium-sized merchants struggle to acquire it [18–20].

Other studies examine traffic distribution from a platform-governance perspective: firms manage resource allocation through search rules, recommendation algorithms, and content-review mechanisms, which shape both merchant behavior and market structure [21]. While the “digital divide” caused by algorithms is well-documented, existing models are primarily grounded in mature urban markets. For rural e-commerce actors who lack brand equity and data analytical capabilities [22], the specific pathways to “traffic inclusion” and resource integration remain insufficiently theorized.

1.1.3. Research on Value Co-Creation and User Participation

Value co-creation theory provides an important theoretical framework for understanding multi-agent interactions within digital platform ecosystems. In digital platform environments, consumers are not only purchasers of products and services but also important participants in value creation. With the rise of social media and live-streaming e-commerce, user participation has become a major source of platform value. Research indicates that in e-commerce platforms, user behaviors such as comments, social interactions, and content sharing can significantly enhance platform activity and brand influence, thereby promoting the development of platform ecosystems [12,23]. Particularly in live-streaming e-commerce models, interactive relationships among anchors, consumers, and brand owners can form highly participatory value co-creation mechanisms. Drawing on dramaturgical and consumption-value theory, Lin et al. (2024) conducted an ethnographic study of how live-streaming sessions guide value co-creation between streamers and audiences, revealing a typology of practices that foster value creation [20]. Additionally, user participation in digital platforms can promote product innovation and brand communication. Consumers' interactive behaviors in platform communities, for instance, help enterprises better understand market demands and enhance product innovation. In their analysis of consumer engagement in live-streaming commerce, Zhang et al. (2024) examined the value co-creation mechanisms and incentive structures that sustain user participation and shape platform performance [21]. From a supply-chain perspective, Li and Yang (2024) deconstructed value co-creation mechanisms, showing how sustained interaction among live-streaming platforms, upstream manufacturers, MCN agencies, and consumers yields multi-party benefits [19]. In the rural e-commerce context, new marketing models such as live-streaming commerce, short video marketing, and social media communication provide new development paths for agricultural product brand communication and market expansion. Based on fieldwork in Z County, central China, Chen et al. (2024) found that integrating social media with e-commerce enables value matching through rich user data, creating new pathways for farmers to build value associations with platforms, users, and suppliers and thus strengthening e-commerce's contribution to rural revitalization [22]. Current research predominantly focuses on large brands or urban communities. How rural entities leverage “rural content” and user participation to trigger algorithmic rewards and acquire traffic lacks a systematic structural explanation.

1.1.4. Research on Complex-Systems Structural Modeling Methods

Integrating rural e-commerce traffic resources involves complex interactions among many factors and therefore calls for complex-systems modeling. In management science and information systems, Structural Equation Modeling (SEM), DEMATEL, and Interpretive Structural Modeling (ISM) are widely used. Structural equation modeling is primarily used to analyze causal relationships and path structures between variables, but its premise

requires researchers to construct clear theoretical models in advance. However, in research on rural e-commerce traffic resource integration, influencing factors often exhibit strong uncertainty and complexity, making it difficult to reveal the underlying driving factors in the system by relying solely on SEM methods [23].

DEMATEL, by contrast, identifies causal relationships among factors and their degrees of influence, surfacing the key drivers in complex systems. ISM converts these relationships into hierarchical structures through matrix operations, exposing core driving factors and transmission paths [24]. Combined DEMATEL–ISM methods have been widely applied to supply-chain management, innovation ecosystems, and digital platform governance [25]. Applying DEMATEL–ISM to rural e-commerce traffic integration therefore helps reveal its structural mechanisms [26]. While these methods are well-suited for analyzing the complexity of traffic integration, their application in rural e-commerce is nascent. Furthermore, the identification of system factors in existing DEMATEL-ISM studies relies heavily on subjective expert judgment, which may lack objectivity in rapidly changing digital environments.

1.1.5. Research Progress on Text Identification Based on Large Language Models and Semantic Embeddings

Large Language Models (LLMs) have gradually become important tools for text analysis. Compared with traditional statistical text analysis methods, large language models can capture semantic relationships and contextual information in texts through deep learning algorithms, thereby improving the accuracy of text analysis. They are now widely applied across management, information systems, and the social sciences. For example, Zhang and Pope (2025) showed that LLMs with ReAct prompting can automate policy-evidence collection—successfully gathering evidence in 45 of 50 experiments using GPT-4o—supporting evidence-based policy-making [27]. Liu and Fu (2024) provided an end-to-end framework for developing and deploying multilingual LLMs in production, combining theoretical foundations with implementation strategies for customer service, search, and machine translation [28]. Additionally, semantic embedding models based on BERT and BERTopic have been used for technology trend identification, industrial innovation analysis, and digital platform research. Lam et al. (2024) introduced concept induction, a computational process that derives high-level, human-interpretable concepts from unstructured text, and showed that their LLoM algorithm outperforms BERTopic in concept quality and data coverage [29]. Grootendorst (2022) originally proposed BERTopic, a neural topic-modeling technique that uses class-based TF-IDF to generate coherent topics by clustering document embeddings; it has since been widely adopted for semantic text analysis [30]. Navarretta and Hansen (2023) applied BERTopic to analyze how Danish political parties addressed energy and environment in their electoral manifestos and parliamentary debates, demonstrating the method’s utility in political discourse analysis [31]. Currently, LLMs are primarily utilized for classification or sentiment analysis. Integrating their semantic-identification capabilities with complex-systems modeling (DEMATEL–ISM) offers a more objective, data-driven foundation for factor extraction in management research.

1.2. Literature Review Summary

The review above identifies five interconnected domains that frame traffic resource integration in rural e-commerce and reveal critical theoretical and methodological gaps. First, while research on rural e-commerce development has shifted from macro-infrastructure to micro-level organizational capabilities, existing studies predominantly treat income or sales as the primary outcome. This leaves the strategic process of how rural actors—who often lack traditional competitive advantages—acquire and integrate platform traffic as a core digital asset largely unexamined. This study addresses this gap by shifting the

analytical focus from performance results to the resource integration processes underlying traffic acquisition. Second, although the mechanisms of platform traffic distribution are well-documented in general e-commerce, they remain grounded in urban contexts where merchants possess high brand recognition and data capabilities. How resource-constrained rural entities navigate “black-box” algorithms to gain visibility, however, lacks theoretical depth; this study responds by exposing the survival and acquisition tactics unique to rural actors in algorithm-dominated ecosystems. Third, value co-creation theory explains user participation and brand influence well, yet its application remains concentrated in urban or corporate brand communities. How “rural authenticity” can be leveraged to trigger user-driven traffic acquisition under rural revitalization remains a theoretical omission this study seeks to fill. Fourth, complex-systems modeling (e.g., DEMATEL–ISM) is well suited to identifying causal hierarchies, yet its application to the non-linear dynamics of rural traffic is nascent. Most e-commerce research still relies on linear models such as SEM that cannot capture emergent transmission paths; this study therefore uses an integrated DEMATEL–ISM framework to map the deep causal logic of traffic influence. Finally, despite the potential of Large Language Models (LLMs) in text analysis, they are currently siloed in sentiment or classification tasks. This study bridges AI-driven text mining and structural modeling by coupling LLM-based semantic identification with formal system analysis.

To address these gaps, the study identifies key influencing factors through semantic analysis of academic literature, policy documents, and industry reports, then traces their causal relationships and hierarchical structures. In doing so, it advances both the theoretical understanding of rural e-commerce traffic dynamics and the methodology for analyzing complex socio-technical systems in digital platform ecosystems.

2. Factors Influencing Rural E-Commerce Traffic Integration

Based on the framework presented in Figure 1, this study proposes a three-stage research pathway—comprising factor identification, relationship measurement, and hierarchical analysis—to investigate the integration of rural e-commerce traffic resources. The framework follows a progressive analytical process, moving from text-driven factor discovery to causal relationship quantification and structural interpretation.

In the factor identification stage, this study develops an integrated LLM–DEMATEL–ISM framework that extends beyond conventional expert-based approaches. Large language models (LLMs) are employed to analyze multi-source textual data, including academic literature, policy documents, industry reports, and news articles, enabling the extraction and conceptual clustering of potential influencing factors. These candidate factors are subsequently refined through structured validation and expert assessment to ensure their conceptual validity and practical relevance.

In the relationship measurement stage, expert judgments are integrated with the DEMATEL method to quantify causal relationships and evaluate the relative influence intensity among the validated factors. Subsequently, ISM is applied to transform the DEMATEL-derived influence matrix into a hierarchical structural model, thereby revealing the transmission pathways and interdependencies among driving, intermediate, and dependent factors.

The methodological contribution of this study lies not in the standalone application of LLMs, but in the integration of LLM-assisted factor discovery, systematic validation, expert evaluation, and DEMATEL–ISM-based structural analysis into a closed-loop analytical framework. This approach bridges data-driven factor identification with expert-validated system modeling, thereby enhancing the systematicity, rigor, and reproducibility of complex system analysis in rural e-commerce research. The specific procedures are described as follows.

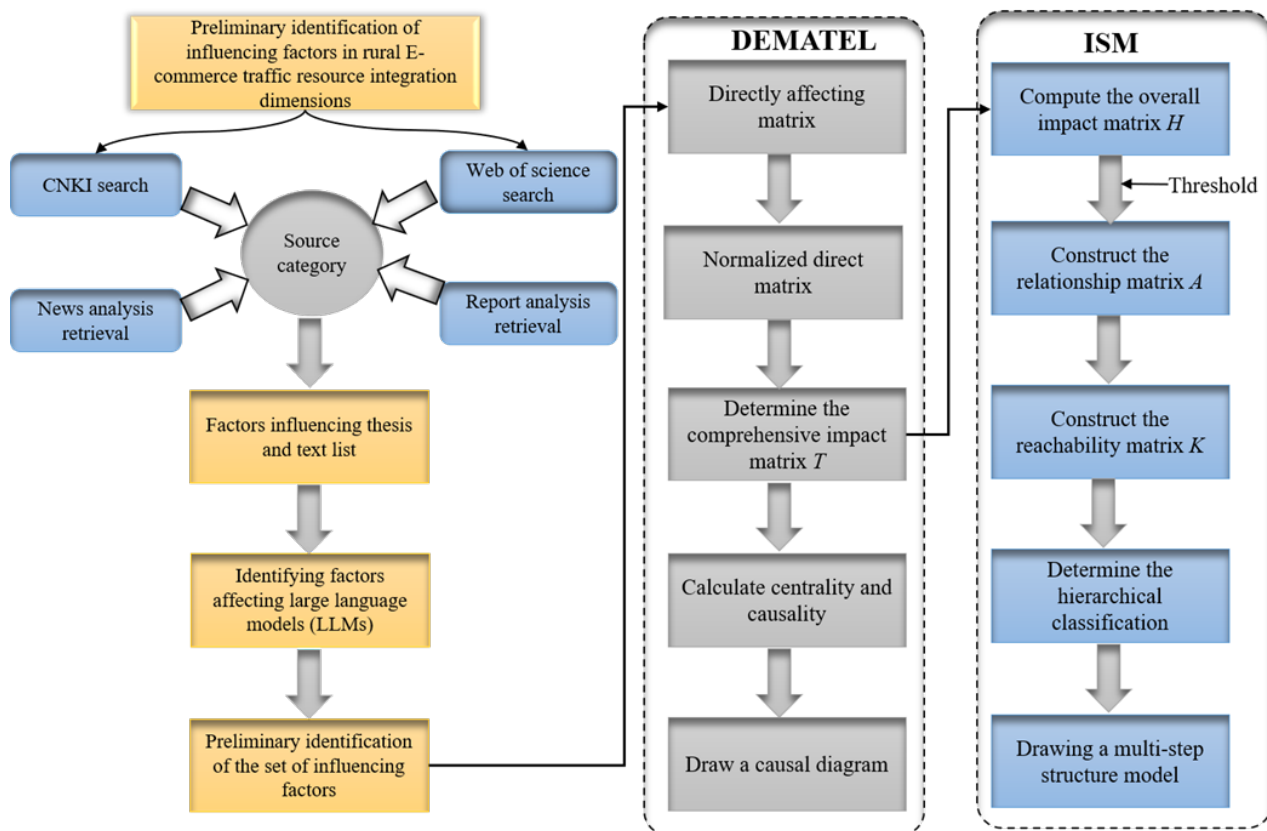


Figure 1. Research framework diagram.

2.1. Phase One: Multi-Source Data Integration and Identification of Influencing Factors

To overcome the limitations of conventional DEMATEL–ISM studies that rely heavily on manual literature reviews and subjective expert elicitation, this study develops an integrated factor identification and validation workflow that combines LLM-assisted semantic extraction, structured screening, expert evaluation, and DEMATEL–ISM modeling. The textual corpus comprised 113 documents collected from academic literature, policy documents, industry reports, authoritative news media, and other relevant sources. After preprocessing, including data cleaning, segmentation, and terminology normalization, the corpus was analyzed using ChatGPT (OpenAI) through the ChatGPT web interface, with ChatGPT-4o accessed from December 2025 to February 2026. Because the ChatGPT interface rather than the OpenAI API was used, API-specific parameters such as top-p and seed were not directly exposed or manually configured. Standardized prompts and identical task instructions were applied across five independent runs ($k = 5$) at a temperature of 0.7. The five runs generated 895, 826, 766, 727, and 894 string-level candidate phrasings, respectively, with an intersection of 136 and a union of 2366 candidates. After semantic consolidation, these outputs were reduced to 413 candidate concepts, whose conceptual boundaries were independently reviewed by two human coders. Detailed prompts, model information, raw outputs, and sampling settings are documented in the Supplementary Materials.

Candidate concepts were subsequently subjected to structured screening based on conceptual equivalence, relevance, representativeness, and conceptual clarity. Expressions with different wording were treated as equivalent only when they represented the same underlying construct and substantive meaning. Redundant, overlapping, ambiguous, or insufficiently relevant concepts were excluded or merged, while the retained concepts were further evaluated by domain experts. This process ensured that the final factor system was not determined solely by LLM outputs but resulted from sequential semantic

verification and expert validation. The resulting 17 factors were then incorporated into the DEMATEL–ISM framework. Expert judgments were used to construct the DEMATEL direct-relation matrix and quantify the influence, centrality, and causal relationships among factors, after which ISM was applied to derive the hierarchical structure and reveal multi-level transmission pathways.

Accordingly, the methodological contribution of the proposed framework lies in establishing a traceable human–machine workflow in which the LLM supports large-scale knowledge extraction and candidate-factor generation, human coders verify semantic consolidation, and domain experts provide substantive validation before DEMATEL–ISM modeling. The complete transition from the source corpus to the initial candidate pool, semantic consolidation, exclusion or merging, expert validation, and the final 17-factor system is illustrated in Figure A1 in Appendix A. The corresponding prompts, model configuration, raw outputs, coding records, and validation materials are provided in the Supplementary Materials to facilitate transparency and reproducibility.

2.2. Phase Two: Causal Strength Measurement and Structural Identification Based on the DEMATEL Model

After identifying the initial influencing factors, this study employs the DEMATEL (Decision Making Trial and Evaluation Laboratory) method to quantify the causal relationships among variables based on expert evaluations. By constructing and analyzing the direct influence matrix, DEMATEL derives key indicators, including influence, dependence, centrality, and causal attributes, thereby identifying critical driving factors and outcome variables within the system. The specific procedures are described as follows.

Step 1: Define the factor set. Assume that the system contains n influencing factors. The factor set is defined as $F = \{F_1, F_2, \dots, F_n\}$, which serves as the common node set for both the DEMATEL direct influence matrix constructed in this phase and the ISM reachability matrix constructed in Section 2.3.

Step 2: Construct the Direct Influence Matrix. Based on the validated factor set, expert evaluations are used to quantify the direct influence relationships among factors. Let x_{ij} represent the influence of factor F_i on F_j , measured on a 5-point scale (0–4), with $x_{ii} = 0$. The aggregated expert scores generate the direct influence matrix $X = (x_{ij})_{n \times n}$, which serves as the basis for subsequent DEMATEL analysis.

Step 3: Construct the Normalized Influence Matrix. To eliminate the scale effects of X and ensure comparability among factors, X is normalized using the row maximum normalization method. Specifically, each element x_{ij} in X is divided by the maximum row sum $s = \max_i \sum_j x_{ij}$, yielding the normalized influence matrix $N = X/s$, in which every entry satisfies $0 \leq n_{ij} < 1$. This bound on N is the analytical prerequisite for the matrix series convergence invoked in Step 4.

Step 4: Calculate the Comprehensive Influence Matrix. Because every entry of N satisfies $0 \leq n_{ij} < 1$, the matrix series $N + N^2 + N^3 + \dots$ converges absolutely, and summing it yields the comprehensive (total) influence matrix $T = N + N^2 + N^3 + \dots = N(I - N)^{-1}$, where I is the $n \times n$ identity matrix. $T = (t_{ij})_{n \times n}$ captures both the direct influence (N) and the indirect influence propagated along all higher-order pathways ($N^2, N^3, \dots$) among factors.

Step 5: Calculate the Influence and Affected Degrees. Based on the comprehensive matrix T , the influence degree and affected degree of each factor are calculated as: $D_i = \sum_j t_{ij}, C_i = \sum_j t_{ji}$ where D_i indicates the total influence exerted by factor F_i , and C_i represents the total influence received by F_i . These two measures characterize the role of each factor in the influence network.

Step 6: Calculate Centrality and Causality. Based on D_i and C_i , the centrality and causality of each factor are calculated as: $M_i = D_i + C_i, R_i = D_i - C_i$ where M_i measures

the overall importance of factor F_i within the system, and R_i indicates its causal role. Factors with $R_i > 0$ are classified as cause factors, whereas those with $R_i < 0$ are identified as effect factors.

Step 7: Construct the Causal Relationship Diagram. The centrality (M_i) and causality (R_i) values are plotted in a Cartesian coordinate system to visualize the structural relationships among factors. The resulting causal relationship diagram distinguishes key driving and effect factors according to their positions, providing the basis for identifying critical variables and supporting the subsequent ISM hierarchical analysis in Section 2.3.

2.3. Phase Three: Hierarchical Structure Division and Path Construction Based on the ISM Model

To reveal the hierarchical structure of the validated factors, this study applies Interpretive Structural Modeling (ISM) to the causal relationships derived from the DEMATEL analysis in Section 2.2. ISM transforms the weighted influence relationships into a hierarchical structure, revealing the structural dependencies and transmission paths among factors. The specific ISM procedures are described as follows.

Step 1: Construct the Adjacency Matrix. Based on the validated factor set $F = \{F_1, \dots, F_n\}$, the adjacency matrix $A = (a_{ij})_{n \times n}$ is constructed by thresholding the comprehensive influence matrix T derived in Section 2.2: $a_{ij} = 1$ indicates a direct relationship from F_i to F_j (that is, $t_{ij} \geq \lambda$), and $a_{ij} = 0$ otherwise, where λ is the threshold value; its determination is described in Section 3.2.1. This transformation converts DEMATEL's weighted influence relationships into the binary structure required for ISM analysis.

Step 2: Calculate the Reachability Matrix. The overall influence matrix is obtained as $S = A + I$, where I is the identity matrix. The reachability matrix is then calculated as: $K = S^r = S^{r+1}$ where r is the smallest integer satisfying the convergence condition. The matrix K represents the reachability relationships among factors and provides the basis for hierarchical classification.

Step 3: Determine the Element Hierarchy. The hierarchical level of each factor is determined based on the reachability matrix K by comparing the following sets: $R(F_i) = \{F_j \in F : k_{ij} = 1\}$, $A(F_i) = \{F_j \in F : k_{ji} = 1\}$, $L(F_i) = R(F_i) \cap A(F_i)$ where $R(F_i)$, $A(F_i)$, and $L(F_i)$ denote the reachable, antecedent, and intersection sets, respectively. These relationships are used to classify factors into different hierarchical levels.

When $R(F_i) = L(F_i)$, factor F_i is assigned to the highest hierarchical level. These identified factors are then removed, and the procedure is repeated until all factors are classified. The resulting hierarchy, combined with the DEMATEL-based centrality and causality indicators, provides the structural foundation for subsequent analysis.

2.4. Validation of LLM-Identified Factors and Their Integration into DEMATEL–ISM

2.4.1. Rationale

The transition from unstructured text to a structured factor set through LLM-assisted factor identification introduces several validity challenges that are not adequately addressed in conventional DEMATEL–ISM studies, where factor identification typically relies on expert judgment or manual literature reviews. These challenges include the stochastic variability of LLM-generated outputs across independent generation rounds, the over-aggregation of causally related concepts into a single factor, and the inadvertent inclusion of system outcomes rather than genuine influencing factors. If left unresolved, these issues may undermine the reliability of factor identification and consequently compromise the validity of subsequent DEMATEL–ISM analysis.

Accordingly, structured validation and expert evaluation are employed to refine and verify the preliminary factor set prior to DEMATEL–ISM modeling, thereby ensuring that only validated factors are incorporated into the influence graph. Because DEMATEL–ISM

operates directly on the validated factor set, this integrated workflow effectively prevents identification errors from propagating through the direct influence matrix X , the total influence matrix T , and the reachability matrix K , thereby enhancing the robustness, transparency, and reproducibility of the overall analytical process. The following section describes the corresponding validation procedure and integration framework.

2.4.2. Three-Tier Validation Protocol

The LLM-identified candidate set F_c is validated through a three-tier protocol combining computational and human-in-the-loop checks (Figure 2). The tiers operate as cumulative gates: a factor must pass all three to enter the final preliminary set F that feeds DEMATEL–ISM.

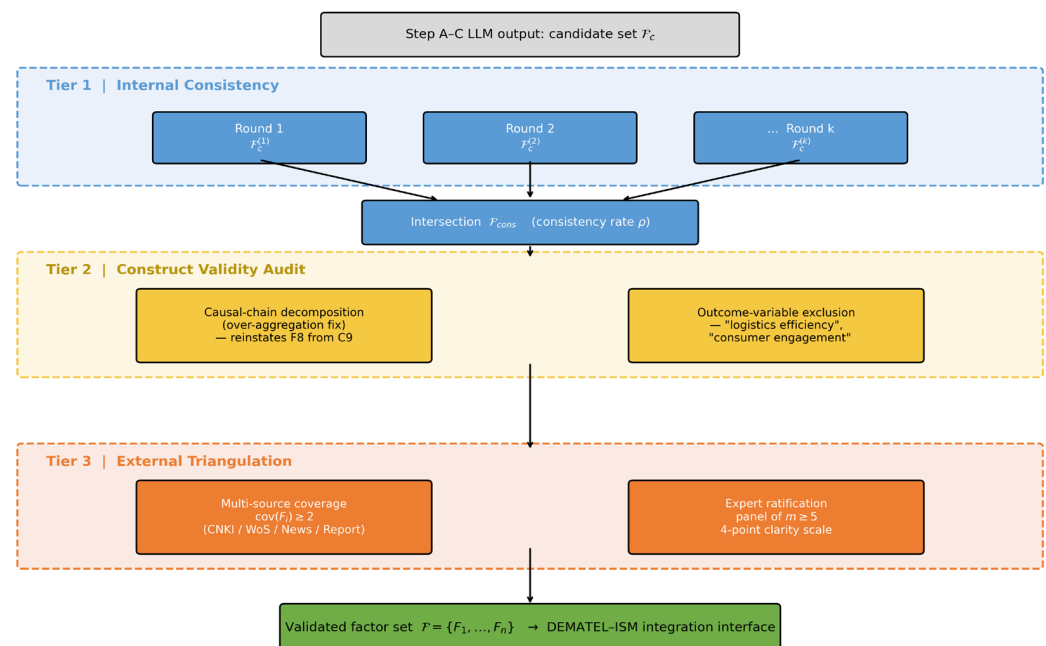


Figure 2. Three-tier validation: Tier 1 enforces multi-round agreement; Tier 2 decomposes over-aggregated chains and excludes outcome variables; Tier 3 requires multi-source coverage and expert ratification. Surviving factors enter set F for DEMATEL–ISM.

Tier 1—Internal consistency (multi-round agreement)

The pipeline is run $k \geq 2$ times on identical inputs with independent sampling; the retained set is the intersection $F_{cons} = \cap_j F_c^{(j)}$, with consistency rate $\rho = \frac{|F_{cons}|}{|\cup_j F_c^{(j)}|}$. To document the evidential value of this check, the identification stage was re-run five times ($k = 5$, temperature 0.7) on the archived 113-document corpus; all prompts, raw outputs, and sampling parameters are archived in the Supplementary Materials. The five runs produced string-level candidate sets of 895, 826, 766, 727, and 894 phrasings, with an intersection of 136 and a union of 2366 ($\rho_{string} = 0.058$): raw outputs are differently worded across runs almost everywhere. After LLM-assisted semantic consolidation of the 2366 phrasings into 413 candidate concepts (cluster boundaries are verified independently by two human coders; the verification worksheet is provided in the Supplementary Materials), the per-run concept sets comprised 296, 255, 278, 252, and 301 concepts, with an intersection of 185 and a union of 413 ($\rho_{sem} = 0.448$). These documented values calibrate the perfect agreement originally reported at the consolidated 17-factor level: consistency at that level of consolidation reflects semantic matching and the near-deterministic extraction settings documented in the prompt registry rather than robustness to stochastic sampling. The

reported ρ is therefore interpreted as an absence of single-pass sampling artifacts, not as evidence of reliability, construct validity, or accuracy, which are assessed in Tiers 2 and 3. A factor that fails to recur across rounds is treated as an artifact of single-pass sampling and is excluded.

Tier 2—Construct validity (boundary audit via iterative prompting)

Each consolidated factor undergoes a structured audit to address two major LLM-related risks. (a) Causal-chain over-aggregation. A consolidation is accepted only when merged elements are semantically equivalent ($=$) rather than causally connected ($\rightarrow$). If a causal chain is identified within a single factor, upstream drivers are retained, intermediate mechanisms are integrated into the most relevant factor, and downstream outcomes are excluded. (b) Outcome-variable exclusion. Variables representing effects rather than causes are removed from F_c to ensure that the retained factors reflect genuine driving mechanisms. These procedures improve causal clarity and conceptual consistency, ensuring that LLM-assisted factor identification provides a reliable input for subsequent DEMATEL–ISM analysis.

Tier 3—External triangulation

Each Tier-2-retained factor undergoes external validation through two criteria. First, multi-source coverage requires that each factor be identified in at least two independent source channels: $cov(F_i) = \sum_s 1[F_i \in F_s] \geq 2$. Factors failing this criterion are subject to additional review. Second, expert ratification is conducted by a panel of $m \geq 5$ experts using a 4-point clarity scale. Factors below the panel median are revised or removed, while unresolved disagreements are addressed through a Delphi-style consensus process. These procedures ensure that the final factor set is empirically supported and conceptually validated for subsequent LLM–DEMATEL–ISM analysis.

2.4.3. Integration Interface to DEMATEL–ISM

The validated factor set $F = \{F_1, \dots, F_n\}$ is transferred to DEMATEL–ISM through an interface standardizing each factor as Dimension | Factor | Definition, preserving dimensional grouping (D1–D4), and recording excluded outcome variables. From F , the forward-direction entries of the direct influence matrix aggregate expert scores as $X_{ij} = \left(\frac{1}{m}\right) \sum_e s_{e,ij}$ (rounded to the 0–4 scale grid), while the reverse-direction entries are derived by LLM-assisted inference as described in Section 3, and $X_{ji} = 0$. The same factor set F is the node set for the ISM reachability matrix K , aligning the DEMATEL and ISM stages.

2.4.4. Transparency and Reproducibility Provisions

Four artifacts are released as Supplementary Materials: (1) Prompt registry—exact prompt templates used at each step (Supplementary Materials). (2) Cached LLM traces—complete LLM input–output pairs at each step in machine-readable JSON. (3) Validation trail—for each retained factor, a structured record of source IDs, consolidation step, multi-round overlap status, dimensional assignment, and expert clarity score. (4) Reference implementation—a Python package (Rural_Ecommerce_Factors) implements the full pipeline with the three-tier validation protocol as composable functions. All of these artifacts, together with the anonymized expert-panel information, the evaluation instrument, the inter-rater reliability results, and the Delphi consensus records described in Section 2.5, have been submitted as Supplementary Materials with this revision and are explicitly linked to the corresponding claims in Sections 2.4 and 2.5.

2.4.5. Demonstrated Contribution of the LLM Identification Stage

The three-tier validation protocol in Section 2.4.2 indicates that the LLM-identified factor set is internally consistent (Tier 1, $\rho = 1.00$, interpreted with the qualifications stated above), conceptually audited (Tier 2), and externally supported (Tier 3). Beyond these checks, the contribution of the LLM identification stage is evaluated from two perspectives. First, single-source restriction analysis shows that no individual channel can recover the complete factor set, indicating that LLM-assisted integration improves cross-source knowledge synthesis. Second, the LLM-identified set was compared with a benchmark factor list compiled independently by the expert panel (Section 2.5) from the same corpus before the LLM outputs were disclosed; all LLM-identified factors were endorsed by the independent coders, and the LLM set additionally recovered emerging platform-related factors (F_{15} , F_{16} , F_{17}) that the manual benchmark initially omitted. The resulting precision, recall, and coverage statistics, together with the coding protocol, are reported in the validation trail provided in the Supplementary Materials. Rather than replacing experts, the proposed workflow establishes a complementary division of labor in which the LLM supports broad and consistent factor discovery while expert validation ensures conceptual accuracy and final ratification. In a documented re-run of the identification stage on the archived corpus, the consolidated concept registry covered 15 of the 17 validated factors by concept-level matching, while the multi-channel integration factor (F_{16}) did not recur, consistent with its originally documented origin in news and industry-report channels; the full registry and run traces are provided in the Supplementary Materials.

The operational characteristics of the LLM-augmented identification pipeline, together with its qualitative comparison with conventional expert-led identification, are documented in the Supplementary Materials, which also contain the prompt registry, cached LLM traces, validation trail, reference implementation, anonymized expert-panel information, evaluation instrument, reliability results, and Delphi consensus records.

2.5. Expert Selection and Evaluation Procedure

Both the DEMATEL matrix X and the Tier-3 ratification rely on a structured expert panel whose validity propagates to all downstream results; its composition, instrument, aggregation, and reliability are specified below.

2.5.1. Panel Composition and Recruitment Criteria

The expert panel comprised six experts holding doctoral degrees and possessing extensive experience in rural e-commerce big data research. Experts were selected based on their academic qualifications, relevant research experience, and practical knowledge of rural e-commerce and digital resource integration. Their complementary expertise supported the validation of the identified factors and the assessment of their interrelationships. The same six experts participated in both factor validation and the subsequent DEMATEL pairwise influence assessment, ensuring consistency between factor identification and structural analysis.

2.5.2. Recruitment Procedure and Ethics

The six experts were recruited through a predefined screening process based on their academic qualifications and relevant research or professional experience in rural e-commerce big data. All six experts completed both the factor validation and DEMATEL evaluation, yielding a 100% completion rate. Following factor identification and validation, the same panel assessed the direct influence relationships among the validated factors using the DEMATEL procedure. All participants provided informed consent, and the study was conducted in accordance with applicable institutional ethical requirements.

2.5.3. Evaluation Instrument

The evaluation instrument comprises two components. First, experts assess factor clarity as part of Tier 3 validation to ensure conceptual consistency before modeling. Second, they conduct pairwise influence assessments to construct the DEMATEL direct influence matrix X , where $X_{ii} = 0$ by definition. Together, these components provide conceptually validated and structurally consistent inputs for the proposed LLM–DEMATEL–ISM framework.

2.5.4. Aggregation Rule, Inter-Rater Reliability, and Consensus Procedure

Panel scores were aggregated by the arithmetic mean, $X_{ij} = \left(\frac{1}{m}\right) \sum_e s_{e,ij}$ with $s_{e,ij} \in \{0, 1, 2, 3, 4\}$ and $X_{ii} = 0$; the forward-direction ratings were collected from the expert panel ($m = 6$), and the reverse-direction intensities were derived by LLM-assisted inference (five runs, per-pair median, temperature-0 control agreement 82.4%; traces in the Supplementary Materials). Inter-rater reliability of the forward ratings was evaluated using Krippendorff's α ($\alpha = 0.10$ across the six raters and 136 pairs), and cells with insufficient agreement were refined through a Delphi-style consensus process.

2.5.5. Procedural Timeline and Transparency Provisions

The expert elicitation and validation procedure followed a structured workflow to ensure traceability and reproducibility. To enhance transparency, three Supplementary Materials are provided: (i) the anonymized expert panel composition, (ii) the evaluation instrument, and (iii) the aggregated influence matrix with reliability indicators and Delphi consensus records.

3. An Empirical Study on Factors Influencing the Integration of Traffic Resources in Rural E-Commerce

The relative importance and interrelationships among factors influencing rural e-commerce traffic resource integration remain unclear. Therefore, this study applies the DEMATEL–ISM model to quantify causal relationships and identify hierarchical structures among influencing factors (see Table 1), providing a theoretical basis for prioritizing key factors and promoting coordinated resource integration.

To obtain reliable evaluation data, six experts with doctoral degrees and extensive experience in rural e-commerce research were invited to assess pairwise influence relationships. Following DEMATEL methodological requirements, each pair was rated on a five-point scale (0 = no influence, 1 = weak, 2 = moderate, 3 = strong, 4 = very strong). Because the elicitation instrument collected the rating for one direction of each pair, the forward-direction scores were aggregated across the six experts by the arithmetic mean (rounded to the nearest integer), while the reverse-direction intensities were derived by LLM-assisted inference: each reverse direction was rated independently in five sampling runs (temperature 0.7) with a temperature-0 control run, and the per-pair median was adopted. All 136 reverse ratings fell within one scale point of the run mean for 100% of pairs, and the temperature-0 control coincided with the median for 82.4% of pairs; the complete inference traces are provided in the Supplementary Materials.

Table 1. Factors influencing the integration of rural e-commerce traffic resources.

Dimension	Influencing Factors	Factor Explanation
Foundational support for trading and performance	F_1 Rural distribution system development	The sophistication of commodity distribution channels and infrastructure enhances supply efficiency.
	F_2 Fresh produce cold-chain transportation capacity	Evaluates cold-chain capacity and delivery timeliness for rural fresh agricultural products.
	F_3 E-commerce last-mile delivery timeliness	Order fulfillment efficiency directly affects consumer experience.
	F_4 E-commerce transaction fulfillment informatization	Assesses digitalization in order management and logistics tracking systems affecting transaction accuracy and transparency.
Carrying capacity of rural e-commerce entities	F_5 Competency of e-commerce professionals	Evaluates practitioners' knowledge, skills, and service capabilities influencing operational efficiency.
	F_6 Scale of e-commerce market entities	Measures the number and scale of market entities reflecting rural e-commerce development potential.
	F_7 E-commerce entity collaboration capability	Evaluates collaboration and resource integration capacity among e-commerce entities.
	F_8 E-commerce entities leading by example	Examines the leadership and demonstration effects of key e-commerce enterprises.
Competitiveness and synergy in agricultural supply	F_9 Agricultural product quality specifications	Reflects agricultural product standardization, quality control, and supply stability.
	F_{10} Distinctive competitiveness of agricultural products	Evaluates product differentiation and regional advantages driving market recognition and premium pricing.
	F_{11} Regional public brand development	The capacity to develop and enhance the value of regional agricultural product brands.
	F_{12} Supply system coordination capability	The capacity to coordinate supply chain resources and improve operational efficiency.
	F_{13} Efficiency of supply–demand matching	Reflects production–consumption alignment and distribution efficiency.
Platform Traffic Allocation and Transmission Mechanism	F_{14} Platform institutional inclusiveness	Evaluates platform accessibility, support policies, and governance mechanisms for rural enterprises.
	F_{15} Traffic algorithm mechanism characteristics	Examines the impact of traffic allocation rules and recommendation algorithms on merchant exposure and competition.
	F_{16} Multi-channel traffic integration	Evaluates multi-platform and online–offline resource integration capacity for market expansion.
	F_{17} Social live-streaming e-commerce	Reflects social media and live-streaming integration in agricultural product marketing innovation.

3.1. Analysis Based on the DEMATEL Model

3.1.1. Computational Process

Following the DEMATEL procedure, Python 3.8 was used to process the data and obtain the comprehensive influence matrix T (Table 2). Based on T , the influence degree (D), affected degree (C), centrality (M), and causality (R) of each factor were calculated (Table 3). Factors with $R > 0$ were classified as cause factors, whereas those with $R < 0$ were classified as effect factors. The corresponding causal relationship diagram (Figure 3) further illustrates the interaction patterns and influence pathways among the factors.

Table 2. Comprehensive impact matrix of factors influencing the integration of rural e-commerce traffic resources.

T	F ₁	F ₂	F ₃	F ₄	F ₅	F ₆	F ₇	F ₈	F ₉	F ₁₀	F ₁₁	F ₁₂	F ₁₃	F ₁₄	F ₁₅	F ₁₆	F ₁₇
F ₁	0.273	0.354	0.334	0.358	0.426	0.363	0.414	0.326	0.338	0.347	0.372	0.419	0.426	0.291	0.321	0.422	0.446
F ₂	0.307	0.284	0.345	0.350	0.397	0.374	0.386	0.318	0.331	0.339	0.363	0.408	0.416	0.304	0.313	0.392	0.435
F ₃	0.337	0.378	0.317	0.403	0.435	0.410	0.445	0.350	0.364	0.374	0.418	0.448	0.475	0.353	0.344	0.452	0.459
F ₄	0.317	0.357	0.357	0.301	0.410	0.387	0.419	0.330	0.362	0.353	0.377	0.403	0.411	0.315	0.345	0.427	0.451
F ₅	0.350	0.350	0.350	0.354	0.365	0.381	0.434	0.361	0.355	0.365	0.409	0.457	0.466	0.344	0.336	0.423	0.486
F ₆	0.346	0.368	0.368	0.373	0.427	0.361	0.476	0.380	0.395	0.425	0.433	0.461	0.468	0.362	0.373	0.467	0.511
F ₇	0.326	0.345	0.345	0.350	0.419	0.395	0.369	0.357	0.351	0.381	0.405	0.452	0.441	0.321	0.332	0.456	0.443
F ₈	0.311	0.330	0.330	0.335	0.402	0.380	0.412	0.283	0.357	0.367	0.391	0.395	0.402	0.309	0.318	0.419	0.443
F ₉	0.301	0.319	0.320	0.325	0.370	0.350	0.379	0.313	0.285	0.374	0.398	0.381	0.388	0.318	0.307	0.405	0.428
F ₁₀	0.306	0.324	0.324	0.329	0.376	0.374	0.405	0.318	0.351	0.300	0.403	0.407	0.395	0.303	0.293	0.411	0.434
F ₁₁	0.301	0.319	0.319	0.323	0.389	0.367	0.398	0.313	0.344	0.354	0.316	0.382	0.407	0.278	0.307	0.385	0.427
F ₁₂	0.339	0.359	0.358	0.344	0.413	0.368	0.420	0.331	0.343	0.352	0.377	0.364	0.452	0.315	0.325	0.428	0.452
F ₁₃	0.332	0.352	0.351	0.337	0.405	0.361	0.412	0.324	0.336	0.345	0.369	0.416	0.363	0.309	0.319	0.419	0.443
F ₁₄	0.214	0.228	0.248	0.251	0.285	0.269	0.292	0.243	0.252	0.259	0.256	0.294	0.299	0.191	0.219	0.297	0.313
F ₁₅	0.234	0.269	0.269	0.273	0.310	0.292	0.316	0.263	0.273	0.260	0.299	0.319	0.325	0.231	0.219	0.342	0.361
F ₁₆	0.303	0.323	0.323	0.346	0.374	0.351	0.400	0.315	0.327	0.335	0.359	0.404	0.411	0.300	0.330	0.348	0.450
F ₁₇	0.304	0.323	0.323	0.327	0.393	0.371	0.383	0.316	0.328	0.337	0.379	0.405	0.413	0.301	0.331	0.409	0.372

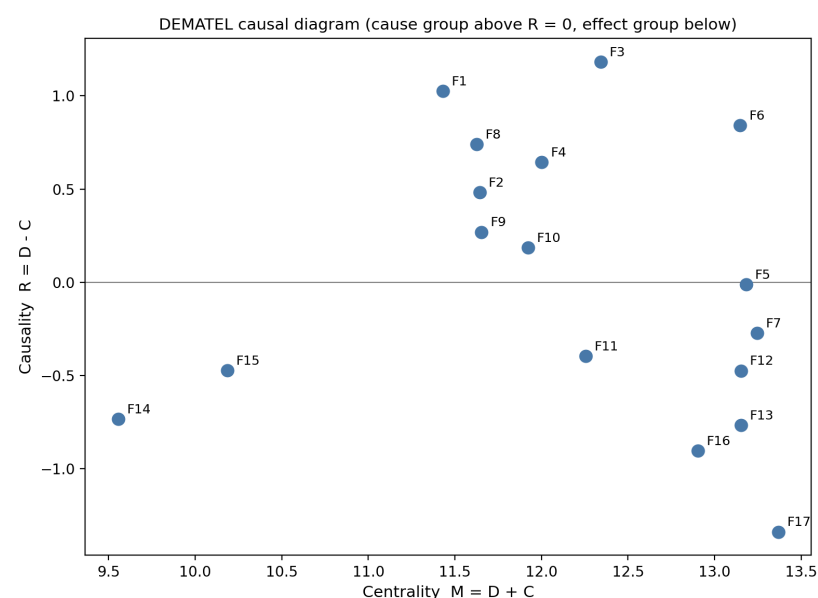


Figure 3. DEMATEL causal relationship diagram, with centrality (M) on the horizontal axis and causality (R) on the vertical axis; factors with $R > 0$ are cause factors, factors with $R < 0$ are effect factors, and factors with $R = 0$ are neutral factors. All 17 validated factors are displayed.

Table 3. Influence, affectedness, centrality, causality, and ranking.

Dimension	Influencing Factors F_i	Influence Degree		Affected Degree		Centrality		Causality		Factor
Foundational support for trading and performance	F_1 Rural distribution system development	6.228	7	5.201	16	11.429	15	1.027	3	cause
	F_2 Fresh produce cold-chain transportation capacity	6.062	10	5.580	12	11.642	13	0.482	10	cause
	F_3 E-commerce last-mile delivery timeliness	6.763	2	5.579	13	12.342	8	1.183	2	cause
	F_4 E-commerce transaction fulfillment informatization	6.323	6	5.679	11	12.002	10	0.643	9	cause
Carrying capacity of rural e-commerce entities	F_5 Competency of e-commerce professionals	6.585	3	6.597	6	13.182	3	−0.012	17	effect
	F_6 Scale of e-commerce market entities	6.995	1	6.153	8	13.148	6	0.842	5	cause
	F_7 E-commerce entity collaboration capability	6.487	4	6.759	5	13.246	2	−0.272	14	effect
	F_8 E-commerce entities leading by example	6.183	9	5.443	14	11.626	14	0.740	7	cause
Competitiveness and synergy in agricultural supply	F_9 Agricultural product quality specifications	5.962	14	5.692	10	11.654	12	0.269	15	cause
	F_{10} Distinctive competitiveness of agricultural products	6.053	11	5.868	9	11.921	11	0.186	16	cause
	F_{11} Regional public brand development	5.930	15	6.325	7	12.256	9	−0.395	13	effect
	F_{12} Supply system coordination capability	6.339	5	6.814	4	13.153	4	−0.475	11	effect
	F_{13} Efficiency of supply–demand matching	6.193	8	6.959	2	13.151	5	−0.766	6	effect
Platform traffic allocation and transmission mechanism	F_{14} Platform institutional inclusiveness	4.410	17	5.144	17	9.554	17	−0.734	8	effect
	F_{15} Traffic algorithm mechanism characteristics	4.856	16	5.330	15	10.186	16	−0.474	12	effect
	F_{16} Multi-channel traffic integration	6.000	13	6.903	3	12.903	7	−0.903	4	effect
	F_{17} Social live-streaming e-commerce	6.014	12	7.353	1	13.367	1	−1.340	1	effect

Figure 4 illustrates the causal network of the rural e-commerce traffic resource integration system, with only relationships exceeding the threshold $\lambda = 0.416$ retained. The network exhibits a core–peripheral structure characterized by dense interactions among several central factors and relatively weak connections for a small number of peripheral factors.

This pattern indicates that traffic resource integration is driven by multiple interconnected mechanisms rather than by a single dominant factor.

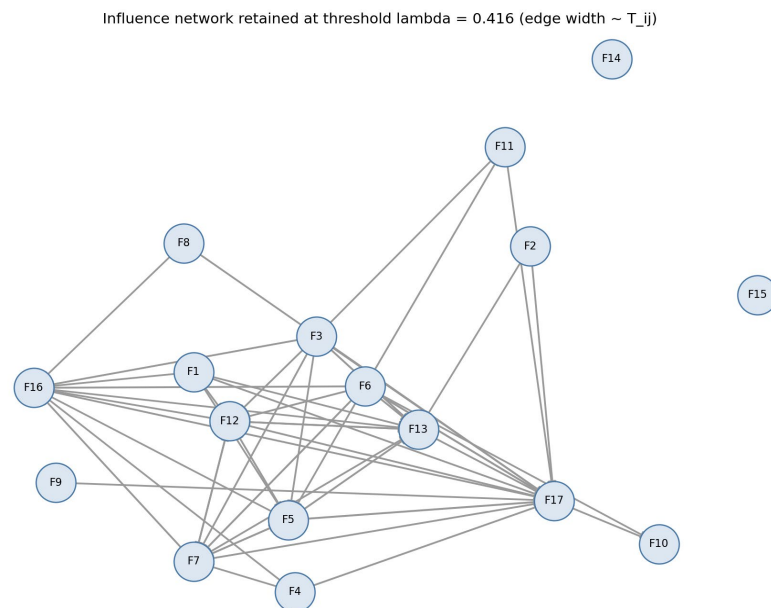


Figure 4. Causal relationship diagram among variables. An edge is drawn when the corresponding element of the comprehensive influence matrix T exceeds the threshold $\lambda = 0.416$ (Equation (2)); weaker links are omitted for clarity. All 17 validated factors are displayed.

F_6 , F_{13} , F_{12} , F_5 , and F_7 constitute the main interaction cluster, with relatively dense connections among capability, supply–demand coordination, and operational factors. In particular, F_6 occupies a prominent position within the network and is closely connected with several other factors, suggesting an important role in transmitting and integrating influences across the system. F_{12} , F_{13} , F_5 and F_7 further form a closely interconnected sub-system, highlighting the joint importance of supply system coordination, supply–demand matching, practitioner capability, and organizational collaboration in supporting traffic resource integration. F_{16} and F_{17} also exhibit extensive connections and can be regarded as important transmission nodes linking multiple parts of the system.

By contrast, F_{14} (platform institutional inclusiveness) and F_{15} (traffic algorithm mechanism characteristics) appear as peripheral nodes because no influence relationship involving these factors exceeds the adopted threshold in the displayed network. This does not imply that they are irrelevant; rather, their direct influences are comparatively weaker than the threshold and are therefore omitted from the visualization for clarity. Similarly, the presence of F_9 and other sparsely connected factors indicates that their strong direct relationships are limited under the selected threshold. Overall, the network combines a densely interconnected core with a relatively sparse periphery, suggesting that improving rural e-commerce traffic resource integration should prioritize the core coordination and capability mechanisms while recognizing the potential supporting role of peripheral platform and algorithmic factors.

3.1.2. Centrality and Causality Analysis

(1) Centrality analysis

Centrality (M -value), reflecting both influence and receptivity, measures the structural importance and hub characteristics of factors within the rural e-commerce system. The distribution of centrality exhibits a clear core–periphery pattern: F_{17} Social Live-Streaming E-commerce ($M = 13.367$), F_7 E-commerce Entity Collaboration Capability ($M = 13.246$),

F_5 E-commerce Practitioner Competency ($M = 13.182$), F_{12} Supply System Coordination ($M = 13.153$), and F_{13} Supply–Demand Matching Efficiency ($M = 13.151$) constitute the core nodes, followed closely by F_6 E-commerce Market Entity Scale ($M = 13.148$) and F_{16} Multi-channel Traffic Integration ($M = 12.903$). These factors maintain extensive bidirectional interactions with other variables and serve as key hubs for resource allocation and information transmission, indicating that rural e-commerce traffic resource integration is jointly sustained by digital marketing mechanisms, organizational capability, and supply-side coordination.

At the local level, the core factors perform distinct structural roles. F_{17} ranks first in centrality and first in affected degree ($C = 7.353$), while its influence degree is moderate (Rank 12), identifying it as the system's principal receiving hub: social live-streaming aggregates influence transmitted from upstream capacity, product, and coordination conditions and converts it into traffic outcomes. F_7 , F_5 , F_{12} , and F_{13} combine high influence with high affectedness, indicating dense two-way coupling between organizational capability and supply-side coordination. By contrast, F_{14} Platform Institutional Inclusiveness (Rank 17) and F_{15} Traffic Algorithm Mechanism (Rank 16) exhibit the lowest centrality, functioning as peripheral, policy-responsive variables rather than core drivers.

Overall, the results demonstrate that rural e-commerce traffic resource integration is sustained by the coupled effects of digital marketing mechanisms (F_{17} , F_{16}), organizational and human-capital capability (F_5 , F_7), and supply-side coordination (F_{12} , F_{13}), while market entity scale (F_6) remains the strongest pure source of influence. High-centrality factors represent the key leverage points for improving system performance and provide priorities for policy intervention and resource allocation.

(2) Causality analysis

Within the DEMATEL framework, the causality index (R) measures the net influence of each factor, calculated as the difference between the influence it exerts and the influence it receives. Combined with the centrality index, the causal diagram distinguishes factors according to both their structural importance and causal roles. Factors with $R > 0$ belong to the cause group and act as system drivers, whereas those with $R < 0$ belong to the effect group and mainly represent outcome or response variables. Classification robustness varies with causality strength: the cause/effect classification remained unchanged across all five inference runs for F_3 , F_6 , F_9 , F_{13} , F_{15} , F_{16} , and F_{17} , whereas factors with causality close to zero (e.g., F_5 , F_{10}) should be interpreted as fragile.

The causality results distinguish a compact cause group from a broader effect group. F_3 (E-commerce Last-mile Fulfillment Timeliness, $R = 1.183$), F_1 (Rural Distribution Systems, $R = 1.027$), F_6 (E-commerce Market Entity Scale, $R = 0.842$), and F_8 (Leading-by-example, $R = 0.740$), together with F_4 , F_2 , F_9 , and F_{10} ($R > 0$), form the cause group: fulfillment, infrastructure, and scale conditions exert net influence on the rest of the system. The classification of F_3 , F_6 , and F_9 was stable across all five inference runs. These factors represent the core operational conditions that generate spillover effects throughout the rural e-commerce system.

In contrast, all remaining factors belong to the effect group. F_{17} ($R = -1.340$), F_{16} ($R = -0.903$), F_{13} ($R = -0.766$), and F_{14} ($R = -0.734$) are the strongest net receivers of influence, followed by F_{12} , F_{15} , F_{11} , F_7 , and F_5 ; no factor is exactly neutral in the aggregated matrix, although F_5 ($R = -0.012$) and F_{10} ($R = 0.186$) lie close to the boundary and their classifications are fragile. At the local level, digital traffic mechanisms (F_{16} , F_{17}) and platform-side governance variables (F_{14} , F_{15}) thus respond to, rather than drive, improvements in upstream fulfillment, infrastructure, and scale conditions.

Overall, the causality analysis indicates that rural e-commerce traffic resource integration is driven primarily by fulfillment timeliness, physical distribution infrastructure, and

market entity scale, while digital marketing mechanisms, platform governance variables, and supply–demand coordination mainly function as receivers and transmitters of these influences. These findings identify the principal leverage points for improving system performance and promoting coordinated rural e-commerce development.

3.2. Analysis of Multi-Level Hierarchical Structural Models

3.2.1. Attainable Matrix

According to the research procedure, a threshold λ is introduced into the comprehensive influence matrix T to remove indicators with relatively weak influence, thereby simplifying the matrix and obtaining the relationship matrix A . The relationship matrix A is then added to the identity matrix I , and Boolean operations are applied through successive matrix multiplications until the results converge, producing the reachability matrix K .

Different threshold values represent different structural relationships among variables. Selecting λ solely based on expert judgment may introduce subjectivity. To improve objectivity, this study determines the threshold using the sum of the mean and standard deviation derived from the statistical distribution of matrix elements. This approach reduces subjective bias and provides a more reliable basis for matrix simplification. The calculation formulas for the threshold λ and the reachability matrix K are presented as follows.

$$\lambda = \alpha + \beta, \lambda \in [0, 1] \quad (1)$$

$$A_{ij} = \begin{cases} 1, & t_{ij} \geq \lambda \\ 0, & t_{ij} < \lambda \end{cases}, i, j = 1, 2, \dots, n \quad (2)$$

Here, α and β denote the mean and standard deviation of all elements in the comprehensive influence matrix T , respectively, while t_{ij} represents the element located in row i and column j of matrix T . Using Python, the mean ($\alpha = 0.361$) and standard deviation ($\beta = 0.055$) of all elements in matrix T are calculated. According to Formula (1), the threshold value is therefore determined as $\lambda = \alpha + \beta = 0.416$. Applying Equations (1) and (2) together with the reachability computation described in Section 2.3 (Step 2), the final reachability matrix K is obtained, as presented in Table 4.

Table 4. Final reachability matrix.

T	F ₁	F ₂	F ₃	F ₄	F ₅	F ₆	F ₇	F ₈	F ₉	F ₁₀	F ₁₁	F ₁₂	F ₁₃	F ₁₄	F ₁₅	F ₁₆	F ₁₇
F ₁	1	0	0	0	1	0	1	0	0	0	0	1	1	0	0	1	1
F ₂	0	1	0	0	1	0	1	0	0	0	0	1	1	0	0	1	1
F ₃	0	0	1	0	1	0	1	0	0	0	1	1	1	0	0	1	1
F ₄	0	0	0	1	1	0	1	0	0	0	0	1	1	0	0	1	1
F ₅	0	0	0	0	1	0	1	0	0	0	0	1	1	0	0	1	1
F ₆	0	0	0	0	1	1	1	0	0	1	1	1	1	0	0	1	1
F ₇	0	0	0	0	1	0	1	0	0	0	0	1	1	0	0	1	1
F ₈	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1
F ₉	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
F ₁₀	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
F ₁₁	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
F ₁₂	0	0	0	0	1	0	1	0	0	0	0	1	1	0	0	1	1
F ₁₃	0	0	0	0	1	0	1	0	0	0	0	1	1	0	0	1	1
F ₁₄	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
F ₁₅	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
F ₁₆	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
F ₁₇	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

3.2.2. Constructing a Multi-Level Hierarchical Structure Model

Based on the reachability matrix, the reachable set, antecedent set, and intersection set are calculated for each factor. These sets are then used to determine the hierarchical levels of the variables, resulting in a hierarchical classification table of factors influencing rural e-commerce traffic resource integration (see Table 5) and the construction of the ISM model diagram (see Figure 5).

Table 5. Hierarchical decomposition.

Hierarchy	Influencing Factors	Hierarchy	Influencing Factors
L1	F_{14}, F_{15}, F_{17}	L3	$F_5, F_7, F_8, F_{12}, F_{13}$
L2	$F_9, F_{10}, F_{11}, F_{16}$	L4	F_1, F_2, F_3, F_4, F_6

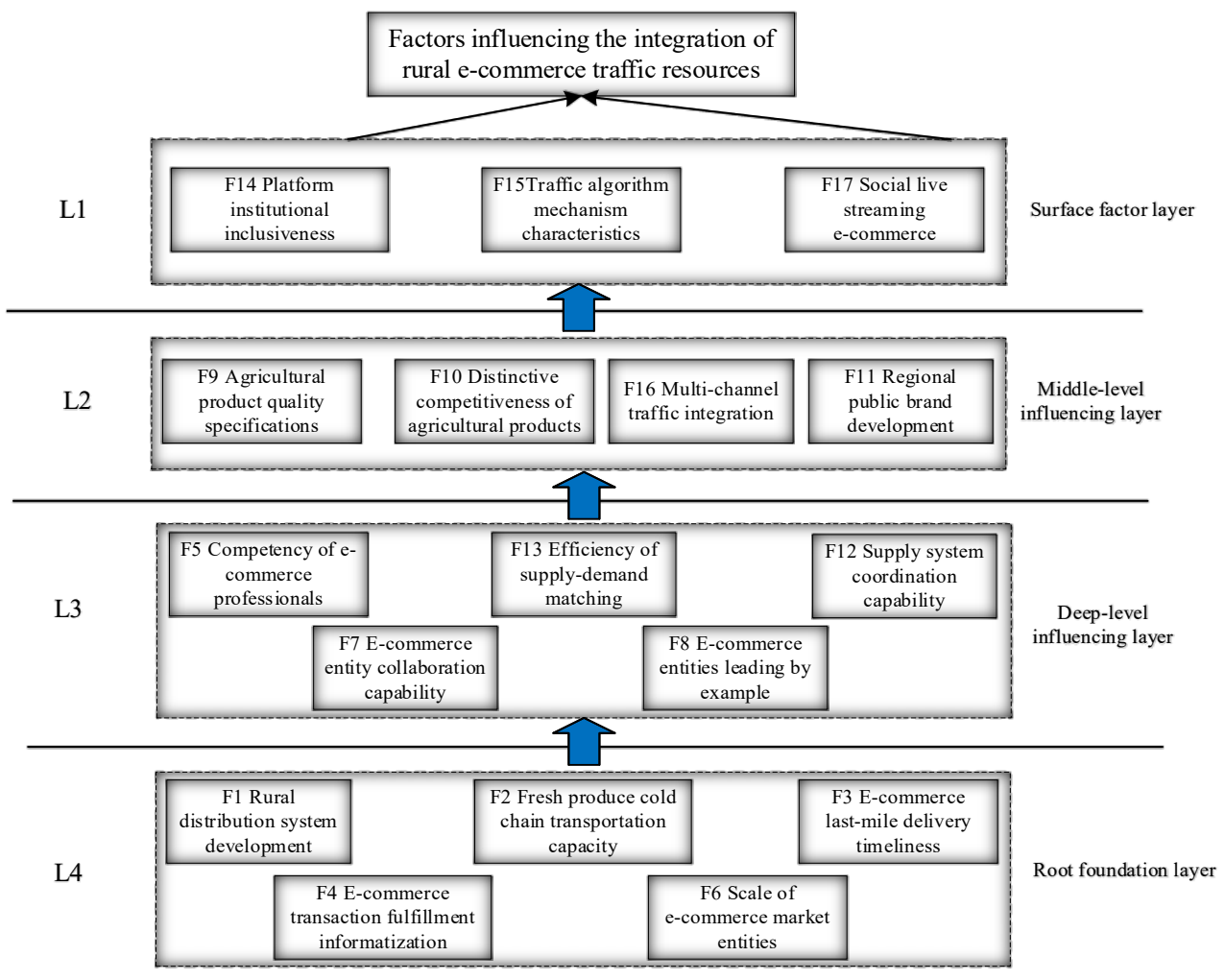


Figure 5. ISM model.

Furthermore, to provide greater transparency regarding the derivation of the hierarchical classification table and the ISM model diagram, Supplementary documents present the adjacency matrix, Boolean multiplication procedure, final reachability matrix, reachability and antecedent sets at each iteration, and the factor-removal sequence used to derive the hierarchical levels. Due to space constraints, these detailed computational procedures are provided in Appendix A to facilitate the verification and reproducibility of the ISM analysis.

3.2.3. ISM Model Hierarchical Analysis

(1) Root foundation layer

The root foundation layer (L4) constitutes the fundamental structural basis of rural e-commerce traffic resource integration and includes F_1 (Rural distribution system development), F_2 (Fresh produce cold-chain transportation capacity), F_3 (E-commerce last-mile delivery timeliness), F_4 (E-commerce transaction fulfillment informatization), and F_6 (Scale of e-commerce market entities). These factors represent the essential infrastructure, logistics, digital fulfillment, and market-capacity conditions upon which higher-level mechanisms depend.

Among them, F_1 and F_2 provide the physical foundation for traffic resource circulation by improving distribution coverage and cold-chain capacity, particularly for agricultural and fresh products. F_3 and F_4 strengthen the operational and digital fulfillment capabilities required to convert traffic resources into efficient transactions, while F_6 reflects the market capacity necessary for absorbing and utilizing traffic resources. As the deepest layer of the ISM hierarchy, these factors exert foundational influence on the factors located in the upper layers. Therefore, policy and managerial interventions should prioritize the improvement of these root conditions, as deficiencies at L4 may constrain the effectiveness of interventions at higher levels.

(2) Deep-level influencing layer

The deep-level influencing layer (L3) comprises F_5 (Competency of e-commerce professionals), F_7 (E-commerce entity collaboration capability), F_8 (E-commerce entities leading by example), and F_{12} (Supply system coordination capability). These factors represent the organizational and capability mechanisms through which the foundational conditions at L4 are transformed into effective system operations.

F_5 strengthens the human-capital base of rural e-commerce, whereas F_7 and F_8 reflect organizational collaboration and demonstration effects among e-commerce entities. F_{12} facilitates coordination across supply-side activities and connects underlying infrastructure with subsequent market and traffic outcomes. Thus, L3 functions as a capability-transmission layer, translating the physical, logistical, digital, and market foundations established at L4 into organizational and operational capabilities.

(3) Middle-level influencing layer

The middle-level layer (L2) includes F_9 (Agricultural product quality specifications), F_{10} (Distinctive competitiveness of agricultural products), F_{11} (Regional public brand development), and F_{16} (Multi-channel traffic integration). These factors represent the market-oriented mechanisms through which organizational capabilities are transformed into product competitiveness and traffic utilization.

F_9 , F_{10} , and F_{11} enhance product quality, differentiation, and market recognition, respectively, while F_{16} facilitates the integration and allocation of traffic across multiple channels. Their position above L3 indicates that market competitiveness and traffic integration are not isolated starting points but are supported by deeper infrastructure, organizational capability, and coordination mechanisms. Strengthening L2 therefore helps convert the capabilities developed at lower layers into tangible market and traffic effects.

(4) Surface factor layer

The surface factor layer (L1) consists of F_{14} (Platform institutional inclusiveness), F_{15} (Traffic algorithm mechanism characteristics), and F_{17} (Social live-streaming e-commerce). These factors are positioned at the highest level of the ISM hierarchy and represent relatively visible manifestations of the underlying system conditions.

F_{14} reflects the inclusiveness of platform institutions, F_{15} captures the characteristics of algorithmic traffic allocation, and F_{17} represents the application of social live-streaming in e-commerce. Their surface-level position suggests that these factors are more directly observable outcomes of the deeper structural and capability conditions rather than isolated root drivers. Accordingly, improvements in platform governance, algorithmic mechanisms, and live-streaming e-commerce should be coordinated with the development of the lower-level foundations rather than treated as standalone interventions.

(5) Overall structure

Overall, the ISM results reveal a four-layer hierarchical structure, progressing from root foundations (L4) through capability transmission (L3) and market-oriented mechanisms (L2) to surface manifestations (L1). The most fundamental layer, L4, comprises distribution infrastructure, cold-chain capacity, last-mile delivery, transaction informatization, and the scale of e-commerce entities, which establish the structural conditions for the operation of the entire system. L3 subsequently transforms these foundations into organizational and coordination capabilities, while L2 converts such capabilities into product competitiveness and multi-channel traffic integration. Finally, L1 reflects the more visible platform and digital-traffic outcomes.

3.3. Analysis of Key Influencing Factors

Combining the DEMATEL results (Tables 2 and 3) with the ISM hierarchy (Figure 5), this section identifies the key determinants of rural e-commerce traffic resource integration from a dual intensity–structure perspective, in which centrality serves as the primary criterion and the influence and affectedness indicators act as complementary validations.

As shown in Table 6, eight variables rank highest by centrality: F_{17} (Social Live-Streaming E-commerce), F_7 (E-commerce Entity Collaboration Capability), F_5 (E-commerce Practitioner Competency), F_{12} (Supply System Coordination Capability), F_{13} (Supply–Demand Matching Efficiency), F_6 (E-commerce Market Entity Scale), F_{16} (Multi-channel Traffic Integration), and F_3 (E-commerce Last-mile Fulfillment Timeliness); six of them (F_3 , F_5 , F_6 , F_7 , F_{12} , F_{13}) also rank within the top eight by influence degree, confirming their roles as core hub nodes within the system.

Table 6. Identification of key influencing factors.

Influencing Factors	Influence Degree Ranking	Affected Degree Ranking	Centrality Ranking	Key Factors
F_{17} Social live-streaming e-commerce	12	1	1	YES
F_7 E-commerce entity collaboration capability	4	5	2	YES
F_5 Competency of e-commerce professionals	3	6	3	YES
F_{12} Supply system coordination capability	5	4	4	YES
F_{13} Efficiency of supply–demand matching	8	2	5	YES
F_6 Scale of e-commerce market entities	1	8	6	YES
F_{16} Multi-channel traffic integration	13	3	7	YES
F_3 E-commerce last-mile delivery timeliness	2	13	8	YES

Among these factors, F_{17} (Social Live-Streaming E-commerce) exhibits the highest centrality ($M = 13.367$) and the highest affected degree, identifying it as the system's principal conversion hub: it absorbs influence from product-side, organizational, and coordination conditions and converts it into traffic outcomes. F_6 (E-commerce Market Entity Scale) exhibits the highest influence degree ($D = 6.995$), retaining its role as the strongest net source of influence and a stable cause factor. F_{16} (Multi-channel Traffic Integration) combines high centrality (Rank 7) with a stable effect classification, indicating a transmission mechanism that couples digital marketing channels with upstream capacity.

Organizational capability constitutes another critical dimension. F_5 (E-commerce Practitioner Competency) and F_7 (E-commerce Entity Collaboration Capability) rank third and second by centrality, respectively, highlighting the importance of human capital and organizational coordination; however, their causality values are close to zero, so their role is better described as high-centrality transmitters than as pure drivers. Meanwhile, F_{12} (Supply System Coordination Capability) and F_{13} (Supply–Demand Matching Efficiency) represent the supply-side coupling between industrial coordination and market matching, and F_3 (E-commerce Last-mile Fulfillment Timeliness) acts as a stable upstream driver connecting logistics efficiency with traffic realization.

Overall, the eight identified key factors—including digital marketing mechanisms, organizational capabilities, supply-side coordination, and fulfillment timeliness—form the core variable cluster of rural e-commerce traffic resource integration. As central nodes within the network, they function as both driving forces and transmission channels that determine system efficiency and stability. Accordingly, optimization strategies should focus on these hub factors by strengthening fulfillment and logistics infrastructure, consolidating market entity scale, deepening organizational coordination, and integrating digital traffic channels. This structural perspective provides a systematic basis for enhancing the efficiency and resilience of rural e-commerce development.

4. Conclusion and Implications

4.1. Theoretical Contributions

The present study offers contributions at three levels: substantive (what the structural analysis reveals about rural e-commerce), methodological (an integrated workflow that combines LLM-assisted factor identification, structured validation, expert evaluation, and DEMATEL–ISM analysis), and epistemological (a structured division of labor between computational intelligence and expert judgment). Each contribution is developed below with reference to the corresponding empirical and methodological findings, and the scope and boundary conditions are discussed in Section 4.1.4.

4.1.1. Substantive Contribution: Reassessing the Causal Position of Platform-Side Factors

The dominant narrative in both academic and policy discourse on rural e-commerce assigns causal priority to platform-side factors: recommendation algorithms, governance inclusiveness, multi-channel integration, and live-streaming are typically framed as upstream levers whose reform would generate downstream economic and developmental gains. The corrected DEMATEL–ISM results reported in Section 3 qualify this framing rather than invert it. The causality indicator R_i (Table 3) classifies all four D4 platform-side factors— F_{14} , F_{15} , F_{16} , and F_{17} —as effect variables ($R_i < 0$, net receivers of influence), with F_{17} ($R = -1.340$) and F_{16} ($R = -0.903$) the strongest net receivers; the classifications of F_{15} , F_{16} , and F_{17} were stable across all five inference runs. The cause group is concentrated in fulfillment and infrastructure conditions (F_3 , F_1 , F_2 , F_4) and market entity scale (F_6). The ISM hierarchy (Figure 5) corroborates this pattern: the deepest layer (L4) comprises F_1 – F_4 and F_6 , whereas F_{14} and F_{15} emerge at the surface layer L1, and F_{17} and F_{16} occupy L1 and

L2 respectively, as outcomes conditioned by upstream conditions rather than engineered independently. Taken together, these results support a measured substantive contribution: platform-side mechanisms are best characterized as high-centrality transmission and reception mechanisms whose development accompanies upstream capacity rather than substituting for it. This reading tempers the platform-centric framing that has dominated recent rural-e-commerce scholarship and indicates that upstream capacity investment is a necessary complement to, rather than a replacement for, platform governance reform.

4.1.2. Methodological Contribution: Validation and Integration Interface for LLM-Identified Factors

The three-tier validation protocol (Section 2.4.2) and the integration interface (Section 2.4.3) together hold LLM-based factor identification to the same reproducibility and construct-validity standards expected of conventional expert-led elicitation. The concrete instance of the F_8 boundary audit (Section 2.4.2, Tier 2a) illustrates why this contribution is non-trivial: the LLM initially consolidated three causally chained concepts into a single node, and the Tier-2 audit detected this over-aggregation, decomposed the chain, and preserved both the originating cause (F_8) and the terminal effect (excluded as a dependent variable) in their correct structural positions. Without this audit, the DEMATEL direct influence matrix would have encoded an incoherent node. This methodological framework is stated in domain-general terms and is directly transferable to other settings in which LLM-identified factor sets feed structural-modeling analyses.

4.1.3. Epistemological Contribution: Hybrid Human–LLM Division of Labor

Rather than treating the LLM as either an autonomous identifier or a mere productivity tool, Section 2.4.5 articulates an explicit division of labor in which the LLM performs the breadth tasks (multi-document reading, cross-language integration, uniform extraction, multi-round consistency) and the expert panel performs the depth tasks (causal-chain adjudication, boundary judgment, final ratification). This articulation matters beyond the present study: the conversation around LLM-augmented research has too often oscillated between uncritical adoption and reflexive skepticism. The hybrid framework locates the LLM within a structured pipeline whose overall reliability rests on the complementarity of its human and computational components, providing a template that other empirical methodologies can adapt.

4.1.4. Generalizability and Boundary Conditions

The four contributions above are bounded, and we state those bounds explicitly. The substantive finding (Section 4.1.1) is grounded in a 2024 snapshot of Chinese rural e-commerce; its generalizability to other national contexts, sectors, or time periods requires empirical verification rather than assumption. The methodological framework (Sections 4.1.2 and 4.1.3) is more general—the protocol does not depend on rural-e-commerce-specific concepts—but its parameter choices ($k \geq 2$, $m \geq 5$, four-channel source structure) should be calibrated to the application. We therefore offer these contributions as transferable templates rather than universal claims, and we release the reference implementation to support replication.

4.2. Conclusions

Amid the integration of digital and platform economies, this study develops an analytical framework combining LLM-assisted factor identification, structured validation, expert evaluation, and DEMATEL–ISM analysis to examine rural e-commerce traffic resource integration. Rather than a simple application of LLMs, the framework addresses the limitations of conventional DEMATEL–ISM studies—particularly their reliance on manually derived

factor sets—by integrating data-driven text mining, expert knowledge, and structural modeling into a unified workflow. Through a stepwise procedure encompassing factor identification, validation, causal analysis, and structural modeling, the study extracts key constructs from diverse textual sources and examines their interrelationships within the system. The findings indicate that the system exhibits a relatively well-defined internal structure, with certain factors playing disproportionately influential roles. In particular, market entity scale, organizational coordination capacity, and digital traffic mechanisms emerge as central nodes linking different subsystems and shaping overall system performance. Factors with stronger causal effects tend to drive system evolution, whereas outcome-oriented indicators primarily reflect the results of these interactions. Among the platform-side mechanisms, F_{14} , F_{15} , F_{16} , and F_{17} all hold effect-variable status, with F_{14} and F_{15} located at the surface layer of the ISM hierarchy.

The hierarchical analysis further reveals that the system operates through a four-layer structure: fulfillment and infrastructure conditions together with market entity scale (F_1 – F_4 , F_6) constitute the deepest foundations (L4); human-capital, organizational, and coordination factors (F_5 , F_7 , F_8 , F_{12} , F_{13}) occupy the intermediate layer (L3); product quality and brand factors (F_9 – F_{11}) together with multi-channel integration (F_{16}) form the second layer (L2); and the platform-side mechanisms (F_{14} , F_{15} , F_{17}) appear at the surface (L1) as observable outcomes. Overall, rural e-commerce traffic resource integration forms a hierarchical, interdependent system in which changes in key driving factors cascade throughout the structure.

4.3. Managerial and Policy Implications

The DEMATEL–ISM results carry direct implications for the design of rural e-commerce development policy and for the strategic decisions of platform operators, rural SMEs, and agricultural suppliers. We organize these implications around the four stakeholder groups most affected by traffic resource integration. Each recommendation is anchored directly to the centrality (M) and causality (R) rankings in Table 3 and to the ISM hierarchy in Figure 5, so that every claim can be traced back to a specific empirical result.

4.3.1. Implications for Government Policy

Three findings bear on the prioritization of public investment in rural e-commerce. First, the ISM hierarchy identifies the deepest layer (L4) as F_1 (rural distribution systems), F_2 (cold-chain capacity), F_3 (fulfillment timeliness), F_4 (transaction informatization), and F_6 (entity scale); public investment in physical distribution, cold-chain, and fulfillment infrastructure therefore propagates upward through every higher layer of the system. Second, the DEMATEL causality analysis (Table 3) classifies F_6 (entity scale, influence-degree rank 1) and F_3 (fulfillment timeliness, $R = 1.183$) as stable cause factors, indicating that policy attention to entity-side scale building and fulfillment conditions produces downstream gains; F_5 (professional competency) and F_7 (collaboration) remain high-centrality factors although their causality is close to zero, so capability-building programs remain justified on centrality grounds. Third, all four D4 platform-side factors (F_{14} – F_{17}) are effect variables, suggesting that platform-side mechanisms function primarily as transmission and reception channels whose development accompanies—rather than substitutes for—upstream capacity building.

4.3.2. Implications for Platform Operators

For platform operators, the centrality rankings highlight F_{17} (social live-streaming e-commerce, $M = 13.367$, rank 1) and F_{16} (multi-channel traffic integration, $M = 12.903$, rank 7) as the highest-centrality D4 factors. Two concrete actions follow: (i) reduce technical and fee barriers for rural merchants on live-streaming and multi-channel tools, since F_{17} and F_{16}

are the principal receiving hubs through which upstream capacity is converted into traffic outcomes; and (ii) refine recommendation algorithms (F_{15} , M -rank 16) and governance rules (F_{14} , M -rank 17) with explicit rural-content weighting, since both are effect variables whose inclusiveness co-varies with, and responds to, merchant-side capability rather than operating as independent levers.

4.3.3. Implications for Rural SMEs and E-Commerce Entities

For rural SMEs, the effect-variable status of F_{14} (platform inclusiveness) means that passive reliance on platform-side policy reform is unlikely to be strategically sufficient. The cause-variable status of F_6 (entity scale, influence-degree rank 1) points to an active scale- and capability-building agenda, while F_5 (professional competency) and F_7 (collaboration) combine top centrality with near-zero causality, indicating that they function as high-centrality transmitters: strengthening them upgrades the system's capacity to absorb and convert traffic. F_8 (leading-by-example) is classified as a cause variable ($R = 0.740$): focal-enterprise leadership feeds upstream capacity building rather than merely emerging from it.

4.3.4. Implications for Agricultural Suppliers

For agricultural suppliers, the supply-side dimension D3 contains one stable cause factor— F_9 (quality specifications, $R = 0.269$, stable across runs)—while F_{10} (distinctive competitiveness, $R = 0.186$) is a fragile cause and F_{11} (regional public brand, $R = -0.395$) is an effect variable at ISM layer L2. The strategic implication is that product standardization remains an upstream lever, whereas regional brand-building functions as a value-receiving mechanism whose returns depend on prior quality and distinctiveness conditions; F_{13} (supply–demand matching, $R = -0.766$, stable) is likewise an outcome of coordination rather than an independent lever.

4.3.5. Synthesis

Across the four stakeholder groups, a consistent pattern emerges: the stable cause factors are fulfillment and infrastructure conditions (F_3, F_1, F_2, F_4) and market entity scale (F_6), which also occupy the deepest ISM layer (L4); the platform-side factors (F_{14} – F_{17}) are uniformly effect variables located at the two uppermost layers; and the organizational and supply-side factors (F_5, F_7, F_{12}, F_{13}) act as high-centrality transmitters in the middle layers. These results support prioritizing fulfillment, infrastructure, and entity-scale investment while treating platform-side mechanisms as complementary, high-centrality transmission channels whose development co-varies with upstream capacity. Table 7 consolidates the stakeholder-specific recommendations and their anchoring evidence.

Each recommendation is traced to a specific empirical result in Table 3 (centrality rank, cause/effect class, run-stability) and Figure 5 (ISM hierarchical role). Leverage tier A = high-leverage intervention targeting a stable cause-class driver, preferably in a deep ISM layer; B = supporting intervention targeting a high-centrality transmitter, a fragile classification, or an effect factor that conditions system-wide performance; C = emergent outcome that should be monitored but not directly engineered. An asterisk (*) marks classifications with causality $|R| < 0.3$ that were unstable across inference runs.

Table 7. Stakeholder-specific actionable recommendations anchored to structural findings.

#	Actionable Recommendation	Stakeholder	Anchored Finding (F-ID, Name)	M Rank	R Class	ISM Role	Leverage Tier
G1	Continue public investment in rural distribution systems and last-mile logistics	Government	F_1 Rural distribution system	15	cause	L4 (deep driver)	A
G2	Expand cold-chain infrastructure programs for fresh agricultural products	Government	F_2 Cold-chain capacity	13	cause	L4 (deep driver)	A
G3	Fund workforce training programs for rural e-commerce practitioners	Government	F_5 Professionals competency	3	effect *	L3 (transmitter)	B
G4	Support MSME incubation and entity-scaling subsidies	Government	F_6 Entity scale	6	cause	L4 (deep driver)	A
G5	Build regional inter-firm coordination platforms (consortia, associations)	Government	F_7 Collaboration capability	2	effect *	L3 (transmitter)	B
P1	Reduce fee and technical barriers for rural merchants on live-streaming channels	Platform	F_{17} Social live-streaming	1	effect	L1 (surface hub)	B
P2	Build and subsidize multi-channel integration tools tailored to rural SMEs	Platform	F_{16} Multi-channel integration	7	effect	L2 (transmitter)	B
P3	Apply rural-content weighting in recommendation algorithms	Platform	F_{15} Algorithm mechanism	16	effect	L1 (surface)	B
P4	Revise platform governance rules to lower entry barriers for rural enterprises	Platform	F_{14} Institutional inclusiveness	17	effect	L1 (surface)	B
S1	Invest in workforce training (operational, marketing, customer service)	SME	F_5 Professionals competency	3	effect *	L3 (transmitter)	B
S2	Join regional entity clusters and supply-chain consortia	SME	F_6, F_7 Entity scale + collaboration	6, 2	cause + effect *	L4/L3	B
S3	Cultivate focal-enterprise leadership visibility as an emergent outcome	SME	F_8 Leading by example	14	cause	L3 (middle)	A
A1	Standardize agricultural product quality specifications and grading	Supplier	F_9 Quality specifications	12	cause	L2 (upper)	A
A2	Participate in provincial-level regional public brand consortia	Supplier	F_{11} Regional public brand	9	effect	L2 (upper)	B
A3	Deploy data-driven supply–demand matching tools	Supplier	F_{13} Supply demand matching	5	effect	L3 (transmitter)	B
A4	Sequence supply system coordination AFTER quality/brand milestones	Supplier	F_{12} Supply system coordination	4	effect	L3 (transmitter)	B

4.4. Research Outlook and Limitations

Although this study develops a structured and practically oriented framework for rural e-commerce traffic resource integration, several limitations remain. First, the proposed framework relies partly on the capabilities of LLMs. Variations in model selection, prompt design, and parameter settings may affect factor extraction and semantic interpretation. Moreover, the knowledge coverage of LLMs is constrained by their training data, which may lead to the omission of emerging, region-specific, or context-dependent factors. Therefore, LLMs should be considered supportive tools for knowledge extraction rather than substitutes for domain expertise.

Second, although structured validation and expert evaluation are incorporated to improve reliability, expert judgment remains essential throughout the factor refinement and causal interpretation processes. Consequently, potential subjectivity may arise when addressing ambiguous or evolving concepts. The proposed framework should thus be viewed as a human–machine collaborative approach that integrates computational intelligence with expert knowledge.

Third, the current framework primarily captures structural relationships at a static stage and does not fully reflect the dynamic evolution of rural e-commerce systems. Future research could incorporate dynamic analytical methods and combine the framework with empirical approaches, such as structural equation modeling or multi-level modeling, to further validate the identified relationships. In addition, regional heterogeneity caused by differences in institutional environments, infrastructure conditions, and economic development levels should be considered.

Finally, integrating network-based indicators, such as centrality and network density, may further enhance the interpretation of DEMATEL–ISM results by identifying key nodes and interaction intensity within the system. Addressing these limitations will improve the methodological robustness and practical applicability of the framework for analyzing rural e-commerce development in complex digital environments.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/jtaer21090318/s1>, Supplementary Materials S1: Figure S1: Rho_distribution_k2; Figure S2: $p_rho1_vs_q$; Figure S3: intersection_vs_k; Table S1: Simulated scoring sheet; Table S2: consolidated_set; Table S3: expert_forward_scores_and_instrument; Table S4: run_factor_sets; Table S5: verify_worksheet_coderA; Table S6: verify_worksheet_coderB; Table S7: X_full_expert_forward_plus_LLM_reverse.

Author Contributions: Conceptualization, Y.H. and J.Z.; Methodology, Y.H. and Q.L.; Software, Q.L.; Validation, Y.H.; Formal analysis, J.Z.; Investigation, Y.H.; Resources, Q.L. and J.Z.; Data curation, Q.L. and J.Z.; Writing—original draft, Y.H.; Writing—review & editing, Q.L. and J.Z.; Visualization, Q.L. and J.Z.; Supervision, Q.L. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset used in this study is derived from primary data collected through expert surveys. The data preprocessing procedures, indicator selection criteria, and experimental code used to generate the empirical results are available from the corresponding author upon reasonable request. To facilitate access and transparency, the relevant data have also been uploaded to a third-party repository: <https://doi.org/10.6084/m9.figshare.33334815>.

Conflicts of Interest: The authors declare no conflict of interests.

Appendix A

For reproducibility, the study should report the LLM model and configuration, access interface and date, language settings, and number of independent runs. The corpus construction process should specify the information sources, search strategy, screening criteria, duplicate removal, and final corpus size. A flowchart should also illustrate the transition from the source corpus to the initial candidates, factor consolidation, and the final 17-factor set. Figure S1 has been added to Appendix A of this paper.

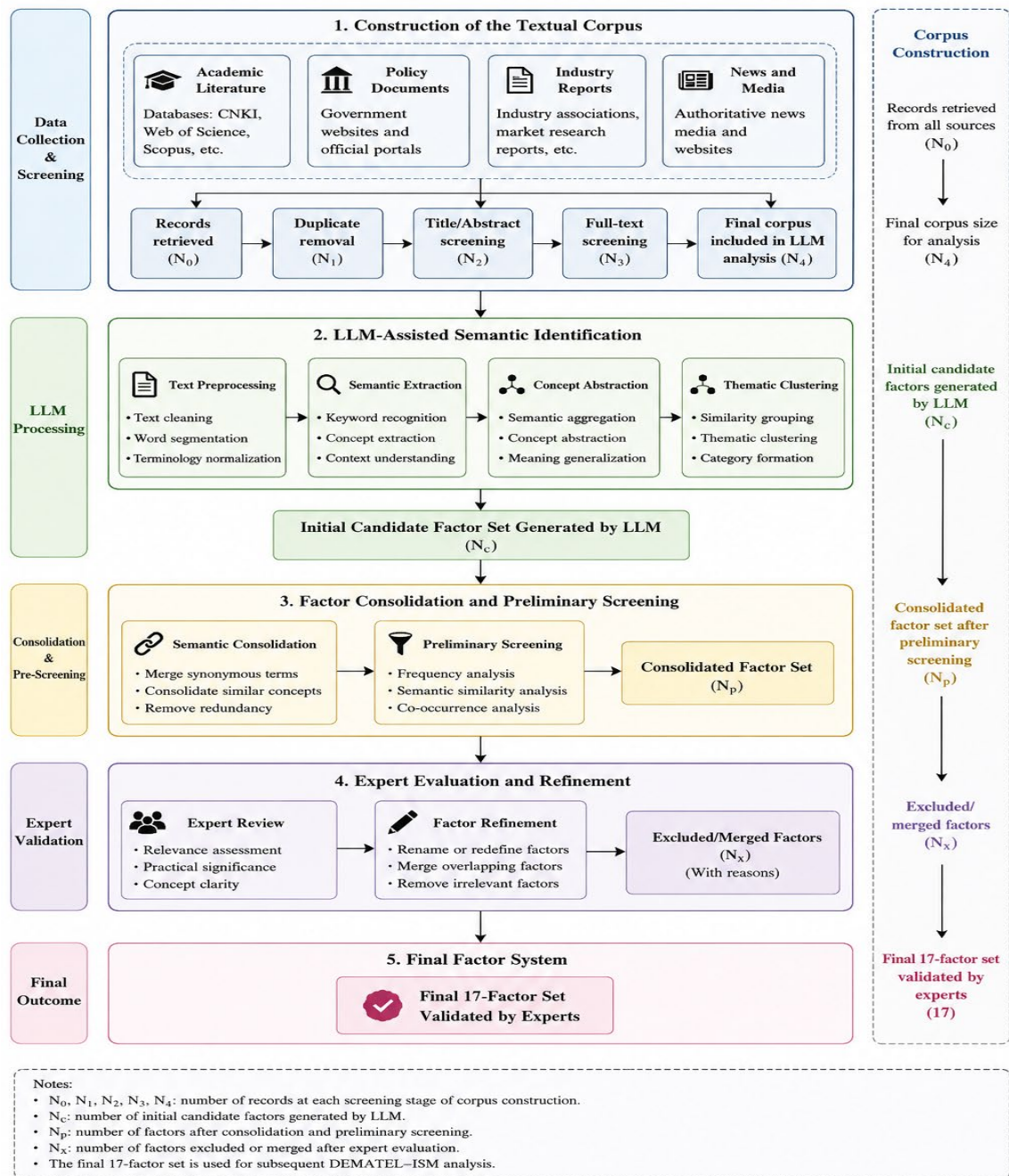


Figure A1. Construction of the textual corpus.

References

- Li, X.; Li, Y.; Chen, Z. Impact of rural E-Commerce participation on farmers' household development resilience: Evidence from 1229 farmers in China. *Agriculture* **2024**, *14*, 692. [CrossRef]
- Liu, H.; Ding, M.; Kan, Y.; Dong, Q. A Study on the Relationship Between Rural E-Commerce Development and Farmers' Income Growth. *Sustainability* **2025**, *17*, 3879. [CrossRef]
- Qiu, H.; Zhang, X.; Feng, M.; Zhang, Z.; Wang, J.; Wang, Z. Exploring the income-increasing benefits of rural e-commerce in China: Implications for the sustainable development of farmers. *Sustainability* **2024**, *16*, 7437. [CrossRef]
- Guan, X.; He, L.; Hu, Z. Impact of rural E-commerce on farmers' income and income gap. *Agriculture* **2024**, *14*, 1689. [CrossRef]
- Zhang, M.; Dong, J.; Zhang, Y. The Impact of Rural E-Commerce Development on Farmers' Income: A Multi-Dimensional Empirical Study. *Res. World Agric. Econ.* **2024**, *5*, 387–402. [CrossRef]
- Cong, H.; Huang, P.; Zou, D.; Zhu, X. Impact of Rural E-commerce on Farmers' Income Increase in Counties of the Yangtze River Delta Region: From the Perspectives of Entrepreneurship and Employment. *Econ. Geogr.* **2024**, *44*, 53–61.

7. Dong, S.; Wang, N.; Fan, C.; Chen, S.; Zhang, L. E-commerce and rural women entrepreneurship—Based on the quasi-natural experiment of “comprehensive demonstration policy” for E-commerce in rural areas. *Econ. Anal. Policy* **2024**, *83*, 749–765. [\[CrossRef\]](#)
8. Li, Q.; Singh, R.; Polpanumas, C.; Fiez, T.; Kumar, N.; Chakrabarti, S. Segment discovery: Enhancing e-commerce targeting. *arXiv* **2024**, arXiv:2409.13847.
9. Liu, S.; Sun, J.; Deng, X.; Wang, H.; Liu, W.; Zhu, C.; Yang, L.T.; Wu, C. Towards platform profit-aware fairness in personalized recommendation. *Int. J. Mach. Learn. Cybern.* **2024**, *15*, 4341–4356. [\[CrossRef\]](#)
10. Wang, Y. Infrastructures for sovereignty: Digital platforms in everyday life. In *Sovereignty Through Practice*; Routledge: Abingdon, UK, 2024; pp. 74–97.
11. Liu, Y.; Zhou, M.; Li, Q. How penetration of e-commerce in rural areas drives consumption in China. *China Agric. Econ. Rev.* **2025**, *17*, 339–357. [\[CrossRef\]](#)
12. Lin, Z.; Souma, T. The Impact of the Pay-What-You-Want Pricing Experience on Customer Behavior: Focus on Willingness to Participate in Value Co-Creation. *Sage Open* **2024**, *14*, 21582440241293304. [\[CrossRef\]](#)
13. Liu, R.; Zhang, L.; Cao, Z.; Mi, J. Influence of gamification affordance on young citizens’ motivation and learning performance toward digital civic education. *J. Chin. Gov.* **2024**, *9*, 244–278. [\[CrossRef\]](#)
14. Zhou, J.; Wang, G.; Diao, Z. Enabling digital transformation of traditional enterprises through industrial internet platforms: A study from an ecosystem perspective. *Eng. Constr. Archit. Manag.* **2026**, *33*, 413–434. [\[CrossRef\]](#)
15. O’Reilly, T.; Strauss, I.; Mazzucato, M. Algorithmic attention rents: A theory of digital platform market power. *Data Policy* **2024**, *6*, e6. [\[CrossRef\]](#)
16. Aird, A.; Farastu, P.; Sun, J.; Stefancová, E.; All, C.; Volda, A.; Mattei, N.; Burke, R. Dynamic fairness-aware recommendation through multi-agent social choice. *ACM Trans. Recomm. Syst.* **2024**, *3*, 1–35. [\[CrossRef\]](#)
17. Huang, X.; Lian, J.; Wang, H.; Liao, H.; Lian, D.; Xie, X. A data-centric multi-objective learning framework for responsible recommendation systems. In *Proceedings of the ACM Web Conference 2024, Singapore, 13–17 May 2024*; Association for Computing Machinery: New York, NY, USA, 2024; pp. 3129–3138.
18. Wang, S.; Zhang, S.; Xu, H.; Gao, T. Developing Market, Product Quality and Traffic Data Game. *Econ. Theory Bus. Manag.* **2024**, *44*, 129.
19. Wang, Y.; Fang, L. Research on the Supply Chain Problems and Optimization of Live Streaming E-commerce. *Front. Bus. Econ. Manag.* **2021**, *2*, 100–104. [\[CrossRef\]](#)
20. Lin, B.; Wang, Y.; Hou, F.; Li, B.; Guan, Z. How live stream session is developed to guide value co-creation between streamer and audience: An ethnographic investigation based on Dramaturgical Theory. In *Proceedings of the 24th International Conference on Electronic Business, Zhuhai, China, 24–28 October 2024*.
21. Zhang, Y.; Xu, Q. Consumer engagement in live streaming commerce: Value co-creation and incentive mechanisms. *J. Retail. Consum. Serv.* **2024**, *81*, 103987. [\[CrossRef\]](#)
22. Chen, Z.; Li, P. Research on E-commerce Boosting Rural Revitalization through Social Communication Mechanism: A Case Study of Z County in Central China. *Journal. Lover* **2024**, 76–78. (In Chinese)
23. Yuan, M.; Hu, H.; Xue, M.; Li, J. Framework for resilience strategies in agricultural supply chain: Assessment in the era of climate change. *Front. Sustain. Food Syst.* **2024**, *8*, 1444910. [\[CrossRef\]](#)
24. Al Amin, M.; Baldacci, R. Blockchain Technology and Industry 5.0 synergy for sustainable development in RMG industries: An ISM and fuzzy DEMATEL approach. *Discov. Sustain.* **2024**, *5*, 464. [\[CrossRef\]](#)
25. Zekhnini, K.; Tarei, P.K.; Benabdellah, A.C.; Oktari, R.S. Analyzing the dyadic relationship between the performance enablers to achieve the humanitarian supply chain 4.0. *Int. J. Disaster Risk Reduct.* **2024**, *115*, 105046. [\[CrossRef\]](#)
26. Panchal, N.; Topre, P.; Kabir, G. Evaluation of Semiconductor Risk Mitigation Strategies in the Electric Vehicle Supply Chain. *Eng. Proc.* **2024**, *76*, 30. [\[CrossRef\]](#)
27. Zhang, Y.; Pope, J. Application of Large Language Models and ReAct prompting in Policy Evidence Collection. In *Proceedings of the ICAART-2025: 17th International Conference on Agents and Artificial Intelligence, Porto, Portugal, 23–25 February 2025*; SciTePress: Setúbal, Portugal, 2025; pp. 967–974.
28. Liu, J.; Fu, B. Responsible multilingual large language models: A survey of development, applications, and societal impact. *arXiv* **2024**, arXiv:2410.17532.
29. Lam, M.S.; Teoh, J.; Landay, J.A.; Heer, J.; Bernstein, M.S. Concept induction: Analyzing unstructured text with high-level concepts using Iloom. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 11–16 May 2024*; Association for Computing Machinery: New York, NY, USA, 2024; pp. 1–28.

30. Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv* **2022**, arXiv:2203.05794.
31. Navarretta, C.; Hansen, D.H. According to BERTopic, what do Danish parties debate on when they address energy and environment? In *Proceedings of the 3rd Workshop on Computational Linguistics for the Political and Social Sciences, Ingolstadt, Germany, 22 September 2023*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2023; pp. 59–68.

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.