

How Relational Dependency Affects the Growth and Development of Global City Networks: A Case Study of Apple iPhone Suppliers (Post-print)

Authors: Liu Qing

Date: 2022-02-11T01:37:10+00:00

Abstract

Current urban network research exhibits a tendency to emphasize structural characteristic measurement while relatively neglecting impact mechanism analysis. Furthermore, in studies on urban network impact mechanisms, there is an overemphasis on theoretical qualitative macro-descriptions or conventional statistical analysis based on independent variables, while micro-level development mechanism research from graph theory and structural perspectives is overlooked. Therefore, from a structural relationship dependency perspective, this study constructs global urban networks of R&D, production, and OEM service types based on Apple supplier data from 2019, and employs exponential random graph models (ERGM) to measure the micro-configurations of urban network growth and development mechanisms. The results indicate: (1) Reciprocity and mediating effects are ubiquitous across the three types of urban networks, profoundly influencing the universal development mechanisms of urban networks. (2) Preferential attachment processes and receiver/sender effects constitute structural mechanisms explaining the development of hierarchical centrality in urban networks, both reflecting a path dependency phenomenon centered on indegree during urban network development. (3) Triadic structures (transitive triads and cyclic triplets) and homophily form the micro-foundation for promoting cluster development and rich-club phenomena in urban networks; the developmental mechanism of homophily in urban networks is primarily manifested in the connection relationships among core cities. (4) Enterprise path dependency and distance are core exogenous factors influencing urban network development. Among these, classified discussions on distance reveal that geographic proximity exerts a pervasive influence on all three types of urban networks. Cognitive proximity demonstrates greater impact and sensitivity on long-distance connections in R&D urban networks; it has a positive effect on medium distances in

production urban networks, while OEM service urban networks exhibit sensitivity to short-distance thresholds. This research holds significant importance for enriching and expanding existing research perspectives on urban network impact mechanisms.

Full Text

Preamble

Current research on urban networks tends to emphasize the measurement of structural characteristics while neglecting the analysis of influence mechanisms. Moreover, existing studies on the influence mechanisms of urban networks often focus on theoretical qualitative macro-descriptions or conventional statistical analyses based on independent variables, paying insufficient attention to micro-level developmental mechanisms from graph theory and structural perspectives. Therefore, from the standpoint of structural relationship dependence, this study constructs R&D-oriented, production-oriented, and OEM service-oriented globalizing city networks based on iPhone supplier data, employing micro-configurations of Exponential Random Graph Models (ERGM) to measure the growth and development mechanisms of these city networks. The results demonstrate that: (1) Reciprocity and intermediary effects are universally present across the three types of urban networks, profoundly influencing the connectivity development mechanism. (2) Preferential attachment processes, together with receiver and sender effects, constitute the core structural mechanisms explaining hierarchical centrality development in urban networks, both reflecting a path dependence phenomenon centered on in-degree during network growth. (3) Triangle structures (transitive triangles and cyclic triads) and homogeneity form the micro-foundations promoting cluster development and rich-club phenomena in urban networks, with homogeneity's influence on network development primarily manifested in linkages between core cities. (4) Enterprise path dependence and distance are core exogenous factors affecting urban network development. Specifically, through classified discussions of distance, geographic proximity exerts ubiquitous influence on all three network types, while cognitive proximity in R&D-oriented networks demonstrates greater impact and sensitivity on long-distance connections. In production-oriented networks, cognitive proximity positively influences medium distances, whereas OEM service-oriented networks show sensitivity to short-distance thresholds. This research significantly enriches and expands existing perspectives on urban network influence mechanisms.

Keywords: globalizing city networks; iPhone; relationship dependence; exponential random graph model

1. Introduction

Since the 1970s, the global expansion and reorganization of multinational corporations have accelerated economic globalization, gradually establishing a new international division of labor centered on global value chain (GVC) fragmentation. Transnational spatial relationships have become a critical analytical lens for studying contemporary globalization geography [?]. Against this backdrop, enterprise-based urban networks represent the mainstream approach in current urban system research [?]. While numerous scholars have contributed to understanding the structural features, dynamics, and developmental patterns of urban networks, a paradigm shift from structuralism to post-structuralism has increasingly emphasized the study of urban network influence mechanisms, constituting an important domain for refining urban network theory.

Recent relational and evolutionary turns in new economic geography have propelled urban network research from structural characteristics toward influence mechanisms. Existing literature primarily adopts three approaches: (1) Theoretical qualitative analysis, drawing from relational and evolutionary perspectives in new economic geography, incorporating concepts such as industrial districts, industrial agglomeration and clusters, location selection patterns, and technological learning and innovation [?]; (2) Conventional statistical analysis, establishing indicator systems from perspectives such as administrative functional hierarchies, markets and costs, competitive advantages, geographic distance, and factor endowments, employing multiple linear regression, Logit regression, negative binomial regression, and Poisson regression models [?], the dominant paradigm for analyzing urban network influencing factors; and (3) Quadratic Assignment Procedure (QAP) correlation and regression analysis. Given that these models test independent variables, some scholars have introduced QAP for non-parametric testing of relational variables when studying trade networks, urban networks, and population mobility networks [?].

However, current research exhibits several relative deficiencies. First, existing literature on urban network influencing factors concentrates on theoretical qualitative analysis or exogenous indicator-based quantitative statistics, with conventional descriptive statistical analysis primarily emphasizing external factors while lacking micro-level configuration studies from structural perspectives on how endogenous network structures affect urban network growth [?, ?]. Second, ERGM-based urban network studies insufficiently explore and classify the effects of path dependence and distance. Research indicates that multinational corporations' location strategies are significantly influenced by existing enterprise networks, with investment behaviors depending on established routines and investment foundations [?]. Furthermore, due to transaction costs, multinationals prefer establishing branches in geographically proximate cities, and social distance proximity also induces inter-city cooperation [?]. Therefore, introducing hypotheses testing path dependence and different distance types in ERGM analysis is necessary.

ERGM, grounded in network dependence theory, posits that relationship interdependence drives network growth and development, realized primarily through triangle structures, reciprocal structures, and star structures [?]. Studies demonstrate that relationship dependence constructively influences urban network development across three dimensions: (1) Relationship dependence affects connection patterns between urban nodes, first manifested in triangle structures—when three cities all connect to a fourth city, these three cities also exhibit strong interconnections [?]. Cross-border production faces challenges of global specialization, shorter product delivery cycles, and more efficient technology alliances, making production linkages between cities effective for resource allocation and complementary advantages, thereby generating synergistic effects [?]. (2) Reciprocal structures constitute a basic configuration for establishing urban connections, reflecting the degree of bidirectional flows between network nodes [?] and measuring network dependence and collaboration. While both exogenous covariates and endogenous structural variables can influence network reciprocity, existing literature primarily focuses on external factors' relationship with reciprocity, with less research on how pure structural effects affect reciprocity under network self-organization. (3) Endogenous structural features also influence the power structure of urban nodes, primarily through star structures that describe node centrality effects [?]. Research finds that both popularity (in-degree) and activity (out-degree) of nodes significantly exist in star structures with few central nodes—in other words, network relationships more easily establish connections among a few “star” nodes with higher network centralization.

In summary, this paper constructs globalizing city networks based on iPhone suppliers, employing ERGM from a relationship dependence perspective to introduce both exogenous covariates and endogenous structural variables into hypothesis testing of urban network development mechanisms. The specific discussion of path dependence and distance' s influence on urban networks significantly enriches and expands existing research perspectives on urban network influence mechanisms. This indicator-deepened discussion also constitutes an important supplement and refinement to structural perspective-based urban network research. Accordingly, this paper addresses the scientific question: How do relationship dependencies affect urban network growth and development through which micro-network configurations? In what aspects are the influences of path dependence and distance manifested? The research proceeds as follows: First, we analyze the structural and local motif characteristics of globalizing city networks, then employ ERGM to measure the development mechanisms of globalizing city networks from endogenous pure structural effects, actor-network effects, and network relational covariates, and analyze how different distance types affect urban network development.

1.1 Construction and Structural Features of Globalizing City Networks

The foundational data for constructing urban networks in this study is iPhone supplier data for 2019, compiled from Apple's supplier directory. We first organized latitude and longitude data for each supplier's headquarters and branches, then queried and compiled business operations and specific components supplied to iPhones through corporate official websites combined with 2019 Apple supplier industry research reports.

The research scale selects prefecture-level city administrative units: all headquarters and branch locations are aggregated to prefecture-level cities and above; for countries with city-affiliated county administrative systems, if branch locations are at the county level, they are aggregated to the county capital or government seat location, ultimately yielding 196 cities. Regarding urban network construction, iPhone components are categorized into four major types: core, specialized, general, and OEM service components, corresponding to three value chain segments: R&D, production, and OEM service (specialized and general components belong to the production segment). The corresponding suppliers are classified as R&D-oriented, production-oriented, and OEM service-oriented. Based on the spatial production layout of headquarters-branches for each category's enterprises across cities, functional connections are established through urban node centrality and connectivity values, constructing R&D-oriented, production-oriented, and OEM service-oriented globalizing city networks (Figure 2). As this paper focuses on analyzing urban network influence mechanisms and is limited by space, please refer to reference [?] for detailed network construction processes, component classification tables, and measurement methods.

The overall characteristics of the three city networks are: (1) The network hierarchical structure is evident, all being polycentric urban networks; (2) A few nodes have numerous functional connections while most nodes have sparse, single connections; (3) A flat network structure exists where a few cities possess both high power and high prestige, with these "star" cities serving as both outward investors and investment destinations, generally holding higher prestige than power. Specific network features are as follows: The R&D-oriented city network exhibits the densest connections, primarily composed of technology center cities, with developed countries' R&D cities geographically proximate to world cities or concentrated as corporate headquarters and R&D institutions, while developing countries host emerging R&D cities oriented toward production support. The production-oriented city network shows the highest connectivity and tends toward uniform structure, involving the largest number of cities, mainly composed of regional central cities in developing countries and a few industrial cities in developed countries. The OEM service-oriented city network demonstrates the most significant hierarchical structure and polarization, with some cities in Taiwan, China as the core, where power and information concentrate, involving the fewest cities primarily located in developing countries.

1.2 Local Motif Characteristics of Globalizing City Networks

Motifs are frequently occurring interconnection subgraphs in real networks, representing the foundation for characterizing group relationships and local generation mechanisms in complex networks, and serving as prerequisites for endogenous structural analysis in ERGM [?]. This study employs Mavisto software to statistically analyze motif relationship frequencies across network types, providing reference for subsequent ERGM analysis. Motif classification and codes follow Davis & Leinhardt's classification system [?], where $\mathbf{G}$ represents a relationship constant term, reciprocity motifs include $\mathbf{G} \leftrightarrow \mathbf{G}$, transitivity motifs include $\mathbf{G} \rightarrow \mathbf{G} \leftarrow \mathbf{G}$ and $\mathbf{G} \rightarrow \mathbf{G} \rightarrow \mathbf{G}$, and clustering motifs include $\mathbf{G} \leftrightarrow \mathbf{G} \leftrightarrow \mathbf{G}$.

From the main motif compositions of the three real networks: (1) In the R&D-oriented city network, the most frequent motifs are $\mathbf{G} \leftrightarrow \mathbf{G}$ and $\mathbf{G} \leftrightarrow \mathbf{G} \leftrightarrow \mathbf{G}$, representing network reciprocity and clustering, indicating significant group-based connections and reciprocal behaviors within clusters. (2) In the production-oriented city network, the most frequent motifs are $\mathbf{G} \leftrightarrow \mathbf{G}$ and $\mathbf{G} \rightarrow \mathbf{G} \rightarrow \mathbf{G}$, containing transitivity (circulation) and reciprocity meanings, indicating high connection accessibility and numerous reciprocal edges in production networks. (3) In the OEM service-oriented city network, the most frequent motifs are $\mathbf{G} \leftrightarrow \mathbf{G}$, $\mathbf{G} \rightarrow \mathbf{G} \rightarrow \mathbf{G}$, and $\mathbf{G} \leftrightarrow \mathbf{G} \leftrightarrow \mathbf{G}$, indicating that reciprocity, transitivity, and clustering are all relatively significant. For clustering, the motif $\mathbf{G} \leftrightarrow \mathbf{G} \leftrightarrow \mathbf{G}$ with intra-regional cooperation characteristics and the cyclic triad structure $\mathbf{G} \rightarrow \mathbf{G} \rightarrow \mathbf{G}$ are most prominent. Overall, reciprocity motifs appear most frequently across globalizing city networks. Additionally, clustering motifs are significant in R&D-oriented and OEM service-oriented networks (represented by $\mathbf{G} \leftrightarrow \mathbf{G} \leftrightarrow \mathbf{G}$), while transitivity motifs are significant in production-oriented networks (primarily $\mathbf{G} \rightarrow \mathbf{G} \rightarrow \mathbf{G}$).

Comparing real networks with random networks (Figure 3): (1) Real networks' average interaction numbers are significantly higher than random networks with identical size and density, indicating that interactivity is a ubiquitous important feature of globalizing city networks. (2) Comparing different motif types, real networks generally outperform random networks in motif types $\mathbf{G} \leftrightarrow \mathbf{G}$, $\mathbf{G} \rightarrow \mathbf{G} \rightarrow \mathbf{G}$, and $\mathbf{G} \leftrightarrow \mathbf{G} \leftrightarrow \mathbf{G}$, while random networks exceed real networks in motif type $\mathbf{G} \rightarrow \mathbf{G} \leftarrow \mathbf{G}$. Compared to random networks of the same scale and density, real networks exhibit more typical interactive, transitive, and cyclic triad relationships. Therefore, endogenous micro-network configurations such as reciprocity, transitivity, and clustering should be considered in ERGM.

2. Methodology

2.1 Exponential Random Graph Model Specification

Exponential Random Graph Models explore city network formation through micro-network configurations. Unlike conventional statistical analyses that observe independent variables, ERGMs are more suitable for observing the conditional dependence of relational variables. ERGM is a network statistical model for relational data that can explain networks at the macro level and hypothesized processes at the micro level [?]. To conceptualize the ERGM, we assume the overall network system as $G(n)$, where n city nodes in the network are represented as $V = \{1, 2, \dots, n\}$, and M represents the set of all possible connections between city nodes. Additionally, G denotes a real network, E is a subset of G representing actual city node connections. Y characterizes elements in M , specifically $y_{ij} = 1$ indicates a multinational corporation's headquarters in city i establishes a branch in city j , otherwise $y_{ij} = 0$. Y is the adjacency matrix of G , and $P(Y = y|\theta)$ represents the probability of y appearing in Y under condition θ . Furthermore, $\theta^T z(y)$ includes three types of network statistics: node attribute covariates, network relational covariates, and network endogenous structural variables, highlighting various interdependence relationships in the network. Therefore, based on the above model assumptions, the basic form is:

$$P(Y = y|\theta) = \frac{1}{k(\theta)} \exp\{\theta^T z(y)\}$$

where Y is the network estimate representing weighted links in the city network; y is the true observed value of Y ; E and $\text{endogeffects}(y)$ represent endogenous structural variables; X and $\text{nodecov}(y)$ represent exogenous city node covariates; G and $\text{edgescov}(y)$ represent exogenous network relational covariates; $k(\theta)$ is a normalization constant ensuring the model is a probability distribution; and $z_n(y)$ represents the network statistics corresponding to $g(y)$, with δ_n as the estimated coefficient for $z_n(y)$ indicating the variable's impact on network formation.

2.2 Definition of Dependent Variables

The dependent variables in this study are obtained by sampling from the 196×196 city network matrices constructed in Section 1.1. First, within these city networks, many cities have relatively sparse and singular connection relationships, with some cities having in-degree and out-degree values of 0, which generates numerous multiplicative factors during model estimation, particularly affecting model convergence with alternating parameters. To ensure robust model estimation, this study overcomes these issues by setting threshold values for sampling in the city networks. Following [?], we employ an upper-tail sampling method, selecting cities with positive out-degree values and capital cities as

samples, ultimately obtaining 67×67 , 85×85 , and 53×53 city sample matrices for R&D-oriented, production-oriented, and OEM service-oriented networks, respectively. Second, since ERGM only analyzes binary network matrices, we dichotomize the three types of sampled city multi-value matrices before model operation, with 0 representing no link and 1 representing an existing link.

2.3 Definition of Explanatory Variables and Research Hypotheses

ERGM examines relationship formation processes through three types of effects: network self-organization processes, node attribute-based processes, and exogenous covariate processes, with network self-organization as the key variable and the remaining two as exogenous control variables. Based on the structural characteristic analysis of the three city networks, this paper selects the following key variables and control variables (Table 1).

2.3.1 Endogenous Structural Effects The micro-configurations of network self-organization processes are selected as endogenous structural effects. This study chooses the following structural parameters as key variables:

- (1) **Arcs** describe the tendency of nodes to form connections. Since networks consist of relationships, arcs as the most basic relational variable count the number of edges, representing the density parameter of city networks.
- (2) **Reciprocity** refers to the degree of bidirectional connections between city network nodes, statistically measured as “city pairs.” Reciprocal cities promote overall network balance through structural equivalence and enhance network interactivity by facilitating inter-city spatial interactions.
- (3) **Preferential attachment** characterizes the tendency of key nodes to contact multiple partners, describing how cities with high prestige (power) further induce prestige (power) agglomeration [?], thereby triggering path dependence in urban network development. Popularity and activity measure node degree distributions (described by k -star statistics), with directed networks distinguishing between popularity (in-degree) and activity (out-degree).
- (4) **Intermediary effect** explains the “middleman” hypothesis by calculating the number of paths connecting two nodes through intermediate nodes, essentially evaluating network accessibility from a connectivity perspective. Connectivity measures the degree to which cities sending relationships simultaneously receive relationships, controlling the correlation between node out-degree and in-degree. This network attribute is simulated through the multiple 2-paths micro-configuration. To consider multiple intermediaries between any node pair, this study uses the alternating k -2-path statistic, which is the weighted sum of k -2-path counts where $k \geq 2$.
- (5) **Clustering effect** is an important structural force affecting network small-group development. Triangle statistics are basic clustering config-

urations, with transitivity and closure mechanisms as their driving forces. Transitive triangles and cyclic triads are jointly interpreted for directed networks with large node-level differences. Transitive triangles assume that two nodes tend to associate mutually when both associate with a third node, promoting triangular relationships [?], simulated in ERGM through alternating k -triangle configurations. Cyclic triads have consistent relationship directions, forming a ternary closure structure as an indirect form of reciprocity.

2.3.2 Actor-Network Effects These effects primarily explore network development from city attribute perspectives, also termed node covariates [?]. The following node covariates are selected:

- (1) **Sender and receiver effects.** Senders and receivers characterize that cities with specific attributes are more likely to send or receive relationships than other cities in the network. Existing research shows that existing industrial bases are important factors in multinational corporations' location strategies [?]-multinationals prefer establishing branches in cities with numerous existing enterprises to obtain agglomeration economies from complete industrial chains or clusters. Therefore, cities with more branches tend to send or receive more relationships.
- (2) **Homophily** refers to nodes with similar attribute features being more likely to connect [?]. Existing social network research finds that network node selection exhibits assortativity, where similar nodes (especially those with high power, prestige, and connectivity values) more easily establish connections, resulting in "rich clubs" or "Matthew effects" [?]. This study measures urban homophily by classifying cities as core cities (node_core) and ordinary cities (node_ordinary) based on degree centrality. Core city homophily counts connections between core cities, while ordinary city homophily counts connections between ordinary cities.
- (3) **City-related attributes (β).** Node endowments and potential affect relationship formation; similarly, cities' resource endowments, economic development levels, and technological capabilities influence city network formation. Existing research identifies GDP, infrastructure, innovation environment, popularity, market size, and market concentration as important factors affecting multinationals' branch location decisions [?, ?, ?]. Therefore, this study introduces six city characteristic variables: β_{1_GDP} (GDP) representing economic development level; $\beta_{2_highway}$ (highway count) representing infrastructure; β_{3_patent} (patent grants) representing innovation environment; $\beta_{4_popularity}$ (Google search results) representing international visibility; β_{5_POP} (population) and $\beta_{6_Herfindahl}$ (Herfindahl index) representing market size and concentration, respectively.

2.3.3 Network Relational Covariates Network relational covariates reflect relationship embeddedness, primarily realized through interactions between

dyadic attribute relationships and statistical data. This section selects the following parameters:

- (1) **Geographic proximity (edg cov₁)**. Existing literature shows that geographic proximity profoundly influences multinational production layout, with inter-city organizational production relationships strongly affected by partner spatial distances [?, ?]. This study calculates Euclidean distances between cities using ArcGIS, representing inter-city shortest distance matrices.
- (2) **Cognitive proximity (edg cov₂)**. This study divides the iPhone value chain into three main division segments: R&D, production, and OEM service. Cities belonging to the same secondary classification under this division are grouped, yielding social distance matrices for the three network types.
- (3) **Openness (edg cov₃)**. Openness reflects a city's production radiation capacity. Research shows that cities with strong radiation capacity, establishing production linkages with numerous cities, affect network structural trajectories [?]. This study uses the number of branch factories a city establishes in another city to represent openness.
- (4) **Path dependence (edg cov₄)**. Enterprise location strategies exhibit path dependence. This indicator uses 2015 iPhone supplier data to construct the 2015 city network, measuring path dependence phenomena in the 2019 city network.

The specific measurement of these variables uses the number of multinational corporation branches located in cities for sender and receiver city attributes. Homophily uses degree centrality for city classification, dividing cities into core and ordinary categories. Due to varying scales and densities across the three network types, different threshold values are adopted: in R&D-oriented networks, cities with degree ≥ 5 are core cities; in production-oriented networks, degree ≥ 8 ; and in OEM service-oriented networks, degree ≥ 4 . City GDP data are from the China City Statistical Yearbook, International Statistical Yearbook, and national economic analysis websites, processed using 2019 average exchange rates. Chinese city highway counts are from www.China-highway.com, while foreign city highway counts are crawled from map.Baidu.com. Patent grant data are from city statistical bulletins and intellectual property offices. City international visibility is measured by Google search result counts. City populations are compiled from Baidu Baike. The Herfindahl index is calculated as the ratio of a city's number of iPhone component product categories to its total number of products (branch factories).

The ERGM in this paper measures factors influencing network growth and development, based on network dependence theory with micro-configurations of triangle structures, reciprocal structures, and star structures. The model construction logic proceeds from endogenous structural effects, actor-network effects, and network relational covariates to establish indicator systems, then

builds regression models based on these micro-configurations and ERGM forms to test relational variables.

3. Results

3.1 ERGM Estimation Results

The statnet software package in **R** estimates ERGM parameters using Markov Chain Monte Carlo Maximum Likelihood Estimation (MCMC MLE). Model complexity and goodness-of-fit are examined through Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) (Table 2). The t -ratio is a convergence statistic; values less than 0.1 indicate proper convergence, while values greater than 0.1 suggest extreme parameter estimates that are unreasonable in model estimation (based on approximate critical values of standard normal variables). The t -ratio values for all three city networks are less than 0.1, indicating appropriately converged parameter estimates. Regarding model goodness-of-fit, AIC values for R&D-oriented, production-oriented, and OEM service-oriented networks are 1,247.23, 1,633.45, and 1,234.56 respectively, with the R&D-oriented network showing the smallest value and thus the best fit. Overall, the AIC values for all three networks remain relatively low, indicating that ERGM estimation results can describe important influencing factors of city networks with acceptable precision.

Arcs. The arcs statistics for all three networks are significantly negative at the 0.1% level. As a constant term, arcs describe the tendency to form connections and network density parameters. The results indicate that globalizing city networks have relatively low overall density, and relationship formation is influenced by specific factors rather than randomly generated.

Reciprocity. Reciprocity is positive and significant across all three networks at the 0.1% level. This demonstrates that reciprocity exerts a universally positive influence in iPhone value chain-based globalizing city networks, playing a fundamental role in network growth and development. Various production factors flow bidirectionally between cities, which prefer to send production factors to cities with which they have established substantial production linkages.

Preferential attachment process. Popularity and activity are represented by gwidegree and gwodegree to characterize network convergence and expansion trends. First, gwidegree is positive across all three networks, indicating that popularity positively promotes globalizing city network development, with clearly investor-preferred cities existing in the networks. Multinationals tend to prioritize these historically popular investment cities when selecting branch locations. However, significance tests show that R&D-oriented and production-oriented networks are significantly positive at the 0.1% and 1% levels respectively, while the OEM service-oriented network is not significant. Regression analysis between branch factories and in-degree centrality also reveals large re-

gression coefficients for R&D-oriented and production-oriented networks, with Pearson correlation coefficients of 0.89 and 0.76 respectively, indicating very strong and strong correlations, with scatter plots concentrated near regression lines. In contrast, the OEM service-oriented network shows a moderate correlation (Pearson coefficient 0.58) with more dispersed scatter distribution. This aligns with ERGM results, indicating that R&D enterprises most strongly prefer locating branches in cities with historically established investment, leveraging accumulated R&D resources and innovation environments, with the highest city selection thresholds. Production-oriented enterprises also show clear investor preferences to utilize established manufacturing foundations and engineering technologies, while OEM service-oriented cities, focusing primarily on low-cost labor advantages, have lower selection thresholds, with high substitutability and transplantability globally, resulting in less significant correlations between branch factories and in-degree centrality.

Regarding activity across the three networks: First, the gwodegree estimate for R&D-oriented networks is negative but not significant, while production-oriented and OEM service-oriented networks show positive but non-significant estimates. Regression tests between branch factories and out-degree (Figure 4) reveal low correlation coefficients (all < 0.3), indicating very weak or no correlation between branch factories and out-degree. Compared to scatter plots of branch factories versus in-degree, these are more dispersed and deviate further from regression lines. This overall suggests that activity is not a significant influencing factor in urban network development—cities sending many relationships do not necessarily tend to send even more. Moreover, in R&D-oriented networks, activity exhibits negative inhibitory effects, indicating that R&D cities prefer network formation based on agglomeration. In production-oriented and OEM service-oriented networks, there are non-significant positive promotion effects, suggesting these networks can form under decentralization forces. Thus, the three network types demonstrate preferential attachment processes with popularity as the core shaping force.

Intermediary effect. This is measured through the multiple 2-paths micro-configuration to assess connectivity's impact. All three networks show positive correlations at the 0.1% significance level, indicating that intermediary effects are universally present in urban network development. Specifically, nodes that both send and receive relationships are ubiquitous in the networks—these nodes with high power and high prestige often become core nodes, guiding the development of network hierarchical structures and enhancing network connectivity accessibility.

Clustering effect. First, the transitive triangle estimates are positive for all three networks, with R&D-oriented and production-oriented networks passing significance tests at the 0.1% and 1% levels respectively, while the OEM service-oriented network is not significant. Regarding cyclic triad closure mechanisms, R&D-oriented and production-oriented networks show positive values at the 0.1% significance level, while the OEM service-oriented network shows negative

non-significant values. This overall indicates that R&D-oriented and production-oriented networks contain numerous triangular closure cooperation relationships, and their larger network densities and scales ensure the construction of ternary relationships. Specifically, cities also tend to connect with their partners' collaborators, forming sub-network level small-group structures where production factor flows, exchanges, and transfers tend to occur within regional clusters, thereby enhancing inter-city cluster cooperation relationships—the primary mechanism for network community structure and cohesive subgroup development.

Sender and receiver effects. First, sender (branch) effect estimates across the three networks are positive but non-significant (none pass significance tests). In contrast, receiver (branch) effect estimates show significant positive impacts (all significant at the 0.1% level). This indicates that sender effects based on branch factories are not significant—cities with more branch factories do not necessarily have higher probabilities of sending relationships to other cities without significant patterns. However, receiver branch factory effects positively promote urban network development, representing path dependence in corporate location selection: enterprises prefer locating branches in cities that already host numerous branch factories. In other words, cities receiving more relationships tend to attract more corporate investment and establish more production linkages.

Homophily. In this study, homophily refers to cities' tendency to connect with similar-featured cities. Across the three networks, homophily based on core cities is significantly positive and passes the 0.1% significance test, while homophily based on ordinary cities is also positive but only significant in production-oriented networks at the 5% level, not significant in R&D-oriented and OEM service-oriented networks. This substantial gap indicates that core city homophily exerts significant positive effects on urban network development. In other words, the tendency for a core city to establish linkages with other core cities is much stronger than with ordinary cities. The development and expansion of the three networks rely more on networking relationships among core cities. This explains the “rich club” phenomenon and negative assortative matching in urban networks—nodes with higher centrality and connectivity values (core cities) tend to network with each other, while peripheral cities with sparse connections have low inter-networking probabilities. Instead, they tend to connect with these high-centrality cities to maximize factor acquisition or resource exchange. Thus, urban nodes expand networks around core city networking relationships, with homophily's influence on network development mechanisms primarily manifested in linkages between core cities.

City-related attributes. In R&D-oriented networks, variables β_{3_patent} , $\beta_{4_popularity}$, and $\beta_{6_Herfindahl}$ are significantly positive (minimum significance at 5%), while other attributes are not significant, indicating that innovation environment, international visibility, and market concentration are main influencing factors in R&D-oriented network development. In production-oriented networks, β_{1_GDP} and $\beta_{2_highway}$ estimates are positive and significant at the 0.1% level, indicating that economic development level and urban infrastructure

significantly and positively affect production-oriented network development. In OEM service-oriented networks, $\beta_{2_highway}$, β_{5_POP} , and $\beta_{6_Herfindahl}$ pass significance tests, indicating that urban infrastructure, market size, and concentration play important roles in OEM service-oriented network development.

Network relational covariates. This study introduces four covariates. First, geographic proximity parameter estimates are significantly positive across all three networks, with R&D-oriented, production-oriented, and OEM service-oriented networks passing significance tests at 0.1%, 0.1%, and 1% levels respectively. In contrast, cognitive proximity shows differentiated positive impacts: R&D-oriented and OEM service-oriented networks pass significance tests at 0.1% and 5% levels respectively, while production-oriented networks are not significant. Results show that geographic proximity significantly and positively affects all three network types, while cognitive proximity's significance varies by network type. Specifically, both geographic and cognitive proximity significantly influence R&D-oriented network development, but cognitive proximity shows the strongest significance. In production-oriented networks, geographic proximity's impact far exceeds cognitive proximity. OEM service-oriented networks show relatively weak impacts from both geographic and cognitive proximity. Regarding the third covariate (openness), estimates are positive but only significantly affect R&D-oriented and production-oriented networks, not OEM service-oriented networks. The fourth covariate, path dependence, shows significantly positive impacts across all three networks at the 0.1% level. Among exogenous covariates, path dependence's influence is the most universal and significant, indicating that city network formation strongly depends on past network foundations. Regression analysis of 2015 city networks' contribution centrality on 2019 networks further confirms this: Pearson correlation coefficients for all three networks range between 0.5-0.8, with scatter distributions concentrated near regression lines, indicating strong correlations between past and present city network development (Figure 5). Indeed, multinationals prefer continuously investing in cities where they have established corporate networks, and current city network development depends on previous industrial foundations.

3.2 Discussion on Distance

This section specifically discusses the relationship between distance and globalizing city networks. First, overall frequency statistics of inter-city distances reveal that city networking patterns universally exhibit clear distance decay patterns, but production-oriented and OEM service-oriented networks show stronger distance decay than R&D-oriented networks. R&D-oriented, production-oriented, and OEM service-oriented networks establish 1,247, 2,156, and 867 relationships within 10,000 km, accounting for 68.3%, 78.9%, and 85.4% of totals respectively. Production-oriented and OEM service-oriented networks decrease with distance, particularly reducing to 156 and 89 relationships in the 24,000-26,000 km range.

In contrast, R&D-oriented network relationship numbers do not consistently decrease with distance, instead showing a growth jump in the 22,000-24,000 km range (Figure 6).

To further investigate underlying patterns, distance is classified into spatial distance and social distance (corresponding to geographic proximity and cognitive proximity above). Boxplots examine how geographic and cognitive proximity affect globalizing city networks under different distance thresholds, revealing significant differences.

First, regarding geographic proximity's impact on the three networks, measuring different geographic distance thresholds shows a negative correlation between geographic distance threshold and influence coefficient. Boxplot fitted lines for all three networks gradually decline with fluctuations. Influence coefficients are highest when geographic distance thresholds reach 2,000 km, but decline power-law significantly when thresholds extend to 4,000 km. Thereafter, as distance thresholds continue expanding, influence coefficients show small-scale fluctuations but overall consistent decreasing trends (Figure 7).

Regarding cognitive proximity's impact, in R&D-oriented networks, cognitive proximity and influence coefficients show a positive relationship—influence coefficients increase with social distance, indicating that in R&D-oriented networks, cognitive proximity has greater impact and higher sensitivity on long-distance cooperative linkages. In production-oriented networks, boxplot fitted lines show a “decrease-increase-decrease” pattern: specifically, when social distance thresholds are 2,000-10,000 km, cognitive proximity's influence decreases; when thresholds expand to 12,000-18,000 km, influence coefficients show a small increase with social distance; when thresholds further expand to 20,000-26,000 km, influence coefficients drop substantially. This overall indicates that cognitive proximity mainly positively influences production-oriented networks in medium distance ranges, with low sensitivity and negative inhibitory effects in short and long distances. In OEM service-oriented networks, boxplot fitted lines continuously decline with fluctuations, indicating that cognitive proximity decreases with social distance increase. This shows that cognitive proximity in OEM service-oriented cities is only sensitive within short-distance thresholds, with insignificant effects on medium and long social distances (Figure 8).

4. Conclusions

This study constructs three types of globalizing city networks based on 2019 iPhone supplier data, focusing on structural relationship dependence to examine urban network growth mechanisms using ERGM, thereby enriching and expanding existing urban network research perspectives. The main conclusions are:

- (1) In local motifs, reciprocity is the most frequent motif across the three

networks, and compared to random networks, the three real networks exhibit more typical interactivity, transitivity, and cyclic triad relationships. Specifically, clustering motifs are significant in R&D-oriented and OEM service-oriented networks, while transitivity motifs are significant in production-oriented networks. Second, ERGM-based structural development mechanism analysis reveals that reciprocity and intermediary effects are universally present across the three networks, profoundly influencing connectivity development mechanisms. Preferential attachment processes and receiver/sender effects constitute core structural mechanisms explaining hierarchical centrality development, both embodying path dependence centered on in-degree during globalizing city network growth. Triangle relationships (transitive triangles and cyclic triads) and homogeneity form the micro-foundations promoting cluster development and rich-club phenomena, with R&D-oriented and production-oriented networks containing numerous triangular closure cooperative relationships and well-developed small-group structures and community networks. Homophily's influence on network development mechanisms is primarily manifested in linkages between core cities.

- (2) Enterprise path dependence and distance are core exogenous factors affecting urban network development. Specifically, classified discussions of distance reveal that geographic proximity ubiquitously influences networking patterns across all three network types, while cognitive proximity in R&D-oriented networks increases with social distance, showing greater impact and sensitivity on long-distance cooperative linkages. In production-oriented networks, cognitive proximity positively influences medium distances, while OEM service-oriented networks are sensitive to short-distance thresholds.

This study finds that preferential attachment and actor effects influence urban network hierarchical centrality development, exhibiting structural dependence phenomena centered on in-degree and receivers, with reciprocity's influence being universal across urban networks—consistent with existing research. However, while some studies suggest that reciprocity and closure mechanisms influence urban network cluster development, this paper finds that reciprocity and intermediary mechanisms also influence connectivity development. Additionally, this study newly explores homophily's role in cluster development, finding that homophily's influence is primarily reflected in core city linkages.

Research limitations include: (1) The discussion of enterprise network path dependence only uses two time-point comparisons (2015 and 2019), lacking continuous longitudinal data to depict a complete evolutionary trajectory. Future research should supplement multiple time-point data to deepen the analysis of enterprise network path dependence. (2) This study explores urban network development mechanisms based on intra-firm value chain linkages, but global value chains also exist in inter-firm relationships. Urban networks formed through different enterprise organizational forms may exhibit different growth mechanisms,

warranting future research from inter-firm value chain dimensions.

References

- [1] Dicken P. Global shift: Reshaping the global economic map in the 21st century[M]. Beijing: The Commercial Press, 2009.
- [2] Ning Yuemin. Spatial organization of enterprise and development of city region[M]. Beijing: Science Press, 2011.
- [3] Chen Xiaofei, Miao Changhong, Pan Shaoqi, et al. Characteristics and construction mechanism of enterprise networks in hub-and-spoke clusters: Empirical evidence from Chery cluster in 2014, China[J]. Geographical Research, 2018, 37(2): 353-365.
- [4] Sheng Kerong, Zhang Hongxia, Zhao Chaoyue. Determinants of the urban spatial network in China: An analysis through the lens of corporate networks within electronic information industry[J]. Geographical Research, 2019, 38(5): 32-46.
- [5] Li Xiande. Spatial structure of the Yangtze River Delta urban network based on the pattern of listed companies network[J]. Progress in Geography, 2014, 33(12): 1587-1600.
- [6] Sheng Kerong, Yang Yu, Sun Wei. Network position and underlying factors of cities in China: A study from corporate networks of the largest 500 public companies[J]. Scientia Geographica Sinica, 2020, 40(5): 68-78.
- [7] Yuan Honglin, Xin Na. Global trade network pattern and influencing factors of advanced manufacturing in China[J]. Economic Geography, 2019, 39(6): 108-117.
- [8] Shao Xuanxuan, Yao Yongling. Spatial network characteristics and influence mechanism of urban agglomerations in the middle reaches of the Yangtze River[J]. Urban Problems, 2019(10): 15-26.
- [9] Sheng Guangyao. Study on the evolution and explanation of inter-provincial population flow network in China[J]. China Population, Resources and Environment, 2018, 28(11): 1-9.
- [10] Ma Wei, Huang Xiaoyan, Cao Xiaoshu. Spatio-temporal pattern of openness of countries along the Belt and Road Initiative and its influencing factors[J]. Arid Land Geography, 2020, 43(5): 207-219.
- [11] Xue Desheng, Zou Xiaohua. The world city network based on the global expansion of Chinese commercial banks and its influencing factors[J]. Acta Geographica Sinica, 2018, 73(6): 989-1001.

- [12] Li Congrong, Xiang Wenqian. Spatial connection intensity of central cities in Guanzhong Plain City Group[J]. Arid Land Geography, 2020, 43(6): 1593-1602.
- [13] Chong Z, Pan S. Understanding the structure and determinants of city network through intra-firm service relationships: The case of Guangdong-Hong Kong-Macao Greater Bay Area[J]. Cities, 2020, 103(5): 1-11.
- [14] Liu X, Derudder B, Liu Y. Regional geographies of intercity corporate networks: The use of exponential random graph models to assess regional network formation[J]. Regional Science, 2015, 94(1): 53-73.
- [15] Ma D, Yu Q, Li J, et al. Innovation diffusion enabler or barrier: An investigation of international patenting based on temporal exponential random graph models[J]. Technology in Society, 2020, 64: 1-12.
- [16] Liu Linqing, Chen Ziruo, Tian Bifei. How does structural dependence affect the formation and evolution of trade network: An empirical analysis based on the Belt and Road[J]. World Economy Studies, 2020, 316(6): 107-121, 138.
- [17] Liu Qing, Yang Yongchun, Jiang Xiaorong, et al. An analysis of the multi-dimensional globalizing city networks based on global value chain: A case study of iPhone' s suppliers[J]. Acta Geographica Sinica, 2021, 76(4): 870-887.
- [18] Boschma R A, Frenken K. Why is economic geography not an evolutionary science? Towards an evolutionary economic geography[J]. Journal of Economic Geography, 2006, 6(3): 273-302.
- [19] Hymer S. The multinational corporation and the law of uneven development: Perspectives on world politics[M]. London: Routledge, 2006.
- [20] Lusher D, Koskinen J, Robins G. Exponential random graph models for social networks: Theory, methods, and applications[M]. Cambridge: Cambridge University Press, 2013.
- [21] Zhang F, Ning Y M, Lou X Y. The evolutionary mechanism of China' s urban network from 1997 to 2015: An analysis of air passenger flows[J]. Cities, 2020, 103(5): 1-10.
- [22] Davis J A, Leinhardt S. The structure of positive relations in small groups[C]//Berger J, Zelditch M, Anderson B, et al. Sociological Theories in Progress. Boston, MA: Houghton Mifflin, 1972: 218-251.

Figures

Source: ChinaXiv –Machine translation. Verify with original.

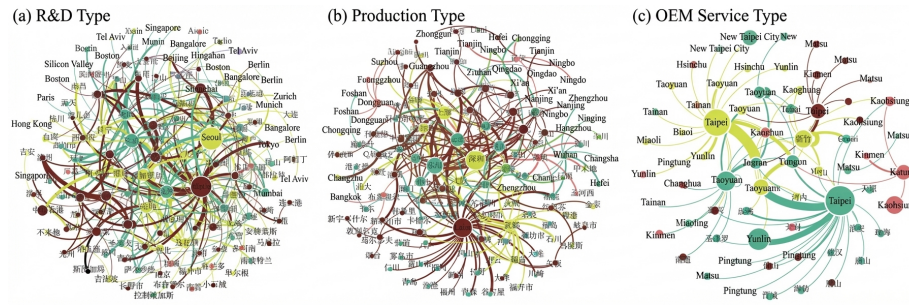


Figure 1: Figure 1

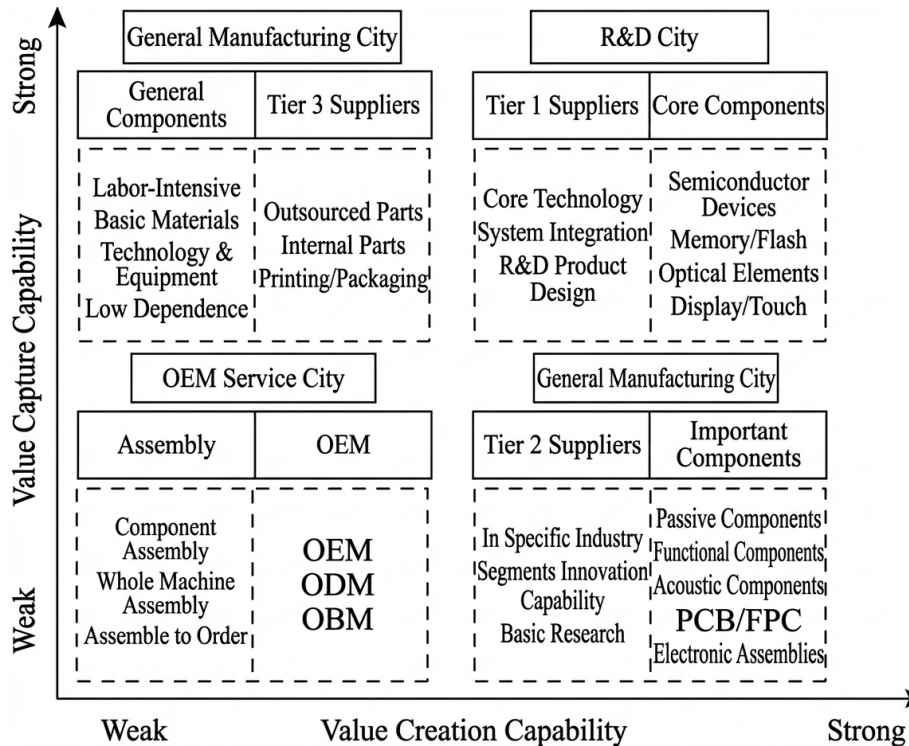


Figure 2: Figure 2

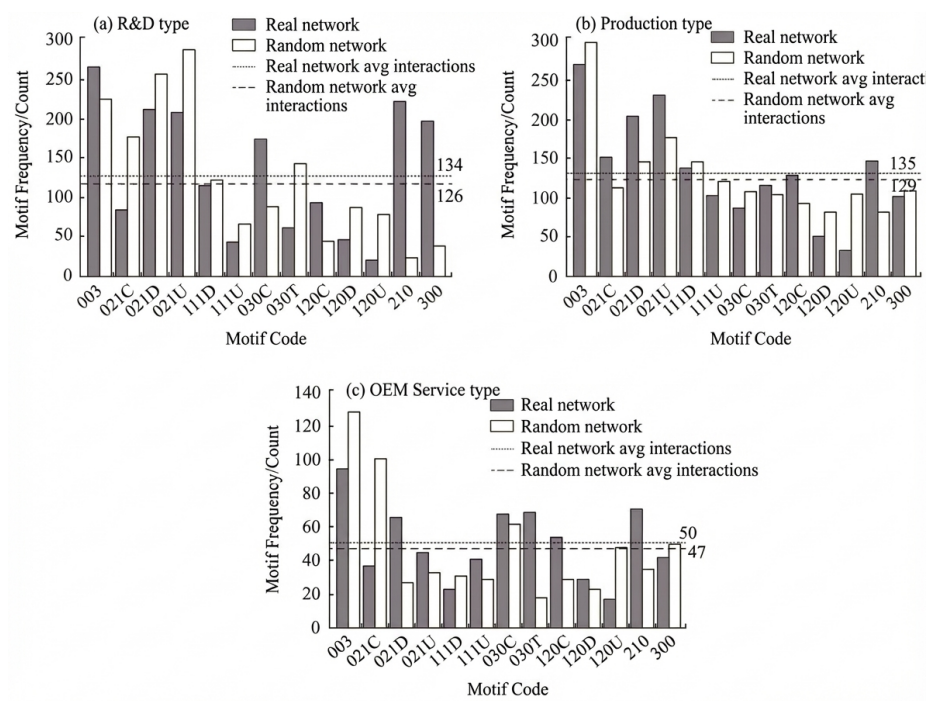


Figure 3: Figure 3

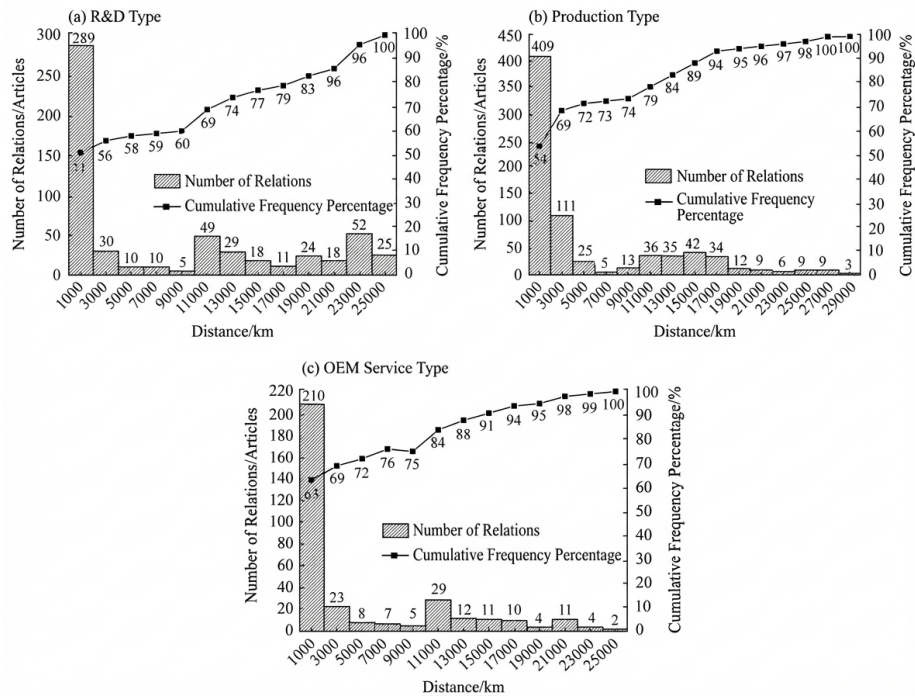


Figure 4: Figure 5

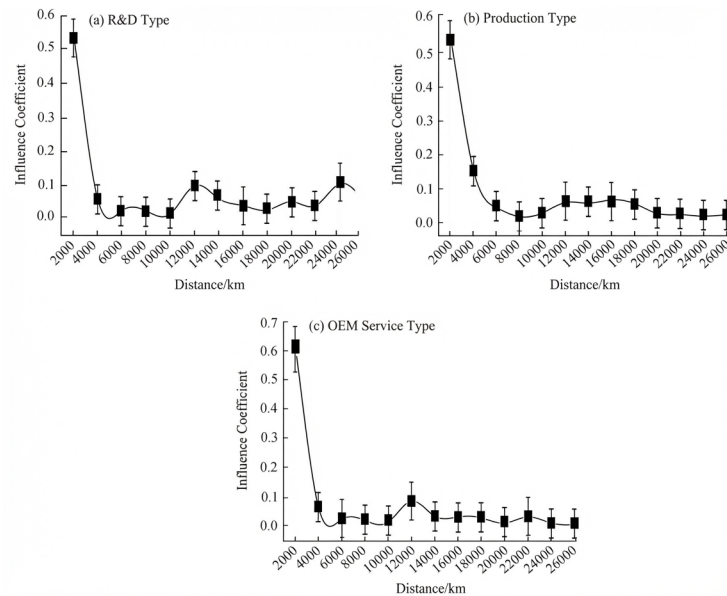


Figure 5: Figure 7