International Journal of Urban Sciences

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rjus20

Examining the evolving structures of intercity knowledge networks: the case of scientific collaboration in China

Liang Dai, Ben Derudder, Zhan Cao & Yufan Ji

To cite this article: Liang Dai, Ben Derudder, Zhan Cao & Yufan Ji (2023) Examining the evolving structures of intercity knowledge networks: the case of scientific collaboration in China, International Journal of Urban Sciences, 27:3, 371-389, DOI: 10.1080/12265934.2022.2042365

To link to this article: https://doi.org/10.1080/12265934.2022.2042365

Published online: 21 Feb 2022.

INTERNATIONAL JOURNAL OF URBAN SCIENCES
2023, VOL. 27, NO. 3, 371–389
https://doi.org/10.1080/12265934.2022.2042365

Routledge
Taylor & Francis Group

Check for updates

Examining the evolving structures of intercity knowledge networks: the case of scientific collaboration in China

Liang Dai [a], Ben Derudder [b,c,d], Zhan Cao [b,e] and Yufan Ji [a]

[a]School of Public Administration, Nanjing University of Finance and Economics, Nanjing, People's Republic of China; [b]Public Governance Institute, KU Leuven, Belgium; [c]Department of Geography, Ghent University, Gent, Belgium; [d]Department of Urban Studies and Regional Development, Nicolaus Copernicus University, Toruń, Poland; [e]College of Architecture and Urban Planning, Tongji University, Shanghai, People's Republic of China

ABSTRACT

Drawing on data on scientific co-publications derived from the Web of Science for the periods 2002–2006 and 2012–2016, we construct and analyse a key element of China's intercity knowledge networks (CIKNs): scientific collaboration networks. Employing network-analytical and exponential random graph modelling techniques, we examine the evolving structures and driving mechanisms underlying these CIKNs. Our results show that the density of the CIKNs has significantly increased over time. CIKN flows are dense in the Southeastern but sparse in the Northwestern part of China, with the Hu Line acting as a clearly visible border. As the dominant knowledge centre, Beijing is involved in scientific collaboration networks throughout the country, with the diamond-shaped structure anchored by Beijing-Shanghai-Guangzhou-Chengdu becoming evident. We find that preferential attachment and transitivity are significant endogenous processes driving scientific collaboration, while a city's administrative level and R&D investment are the strongest exogenous factors. The impact of GDP and geographical proximity is limited, with institutional proximity being the most sizable of the well-known suite of proximity effects.

ARTICLE HISTORY
Received 22 September 2021
Accepted 28 January 2022

KEYWORDS
Intercity knowledge network; scientific collaboration; social network analysis; exponential random graph model; China

Highlights

- China's intercity knowledge networks are built on scientific co-publication data.
- Flows are dense in the Southeastern but sparse in the Northwestern part of China.
- A diamond structure anchored by Beijing-Shanghai-Guangzhou-Chengdu becomes evident.
- Exponential random graph models help reveal the exogenous and endogenous forces.
- Significant endogenous forces include preferential attachment and transitivity.

CONTACT Zhan Cao 1989caozhan@tongji.edu.cn College of Architecture and Urban Planning, Tongji University, Siping Road 1239, Shanghai 200092, People's Republic of China

1 Introduction

Knowledge creation and diffusion among cities have drawn considerable attention in urban and regional science because new knowledge is increasingly regarded as a strategic resource for sustained economic growth of cities and regions. Cities are often argued to be the 'natural' site for the development of interlocking pools of talent, knowledge institutions, and enterprises, thus becoming spatial incubators for knowledge innovation (Florida et al., 2017). With the accelerated complexity of knowledge and knowledge innovation, key actors in this process are pursuing cross-organizational and cross-regional knowledge collaboration, forming an open innovation system (Chesbrough, 2003). From this perspective, an intercity knowledge network is formed by various formal and informal collaborative linkages between the cities in which these actors are embedded. Research focuses on interconnected systems of knowledge flows aggregated at the city level because knowledge collaboration is also a strategic process for cities to achieve the joint complementarity and optimization of innovative resources (van der Wouden and Rigby, 2019). Therefore, knowledge production in a city does not merely depend on processes in its 'local' knowledge base, but also on its embeddedness in wider knowledge networks and accessibility of external knowledge flows.

Data on scientific journal articles, patents, and research and development (R&D) projects are often-used proxies to construct intercity knowledge networks. Co-publications, co-patents, and co-projects are a joint work among co-authors who share ideas, discuss, interact, and jointly produce a qualified output (Liefner and Hennemann, 2011). When the authors are affiliated to organizations in different cities, the knowledge collaboration between them can be envisaged as a proxy for knowledge flows between the cities from which they work. Therefore, these data have served as the input to analyses of how knowledge spillovers, transfers, and collaborations (re)shape urban systems at different scales. For instance, Matthiessen et al. (2010) demonstrated the hierarchical and regional tendencies in the world city system through joint publications between researchers located in different cities. Yao et al. (2020) investigated the spatial patterns and topological features of China's national innovation network based on intercity co-patent data, while Lata et al. (2015) conducted a comparative analysis of intercity knowledge flows in the European Union through the lens of project-based R&D networks, co-patent networks, and co-publication networks. In empirical studies, publications and patents are the most often used data for capturing tacit knowledge exchanges with the former representing scientific knowledge and the latter representing technological knowledge.

Previous studies dealing with intercity knowledge networks in urban and regional science mostly focused on two domains. On the one hand, scholars examined the spatial and topological structures of the knowledge networks, tracing their evolution by means of spatial analysis, social network analysis, and complex network analysis (Andersson et al., 2014; Li and Phelps, 2017; Cao et al., 2021a). Centrality measures, network partition techniques, topological statistics are used to reveal hierarchies, subgroups, and complexity at the node, community, and network levels, respectively. Other topics include the heterogenous spatial distribution of knowledge flows, the degree of knowledge polycentricity, and core–periphery structures. On the other hand, scholars discussed the driving mechanisms underlying these intercity knowledge flows by means of spatial interaction models, negative binomial regression and gravity-type models, and multiple regression quadratic assignment procedures (MRQAP) (Scherngell and Hu, 2011; Gui et al., 2019; Zhang et al., 2020b). These approaches primarily shed light on how city attributes and multidimensional proximity (i.e. geographical, social, cognitive, institutional, and cultural proximity) influence the formation and strength of intercity knowledge flows (Boschma, 2005). However, much of this research has above all focused on exogenous forces such as city attributes and dyadic attributes, thus ignoring endogenous forces. Recent advances in social network research have shown that networks are shaped by both exogenous and endogenous effects (Snijders et al., 2006, 2010). Exogenous effects denote attribute-based factors of nodes and edges, while endogenous effects denote structure-based factors of local network configurations/structures, processes that cannot be captured by conventional (spatial) econometric models.

Drawing on complex-network analytical tools, research has increasingly turned to the analysis of the topology of the network *per se*. Some knowledge networks are shown to be scale-free and small-world networks (Li et al., 2015; Duan et al., 2018), characterized by self-organization processes such as transitivity (Vertes et al., 2012), preferential attachment (Vinciguerra et al., 2010), and reciprocity (Zhang et al., 2020a). Work in this vein mainly focused on statistical descriptions and has therefore rarely reported on the potential influence of endogenous processes on the formation and evolution of intercity knowledge networks.

Against this backdrop, the objective of this paper is to complement this literature by specifying and applying an exponential random graph model (ERGM) to the study of knowledge networks. An ERGM considers both exogenous forces and endogenous forces, and is used here to address two more concrete research questions: (1) What are the spatial and topological structures of intercity knowledge networks and their evolving trajectories in China? (2) How do exogenous and endogenous forces affect the formation and evolution of intercity knowledge networks in China? The research focuses on the case of one of China's intercity knowledge networks (CIKNs), i.e. collaborations between scientists working in different cities. Our analyses draw on co-publication data retrieved from Web of Science (WoS) for the periods 2002–2006 and 2012–2016. The remainder of this article is organized as follows. The second section gives an overview of earlier models and possible determinants of intercity knowledge networks. The third section elaborates our research design, including the construction and characterization of CIKNs, the overview of ERGMs, and the specification of variables. The fourth section presents and discusses the empirical results, followed by conclusions and avenues for future research in the final section.

2 Literature review

2.1 Explanatory models for intercity knowledge networks

Drawing on a range of network-analytical techniques, various approaches have been put forward to model intercity knowledge networks. These approaches range from standard (spatial) econometric models, to permutation-based MRQAP models, and to stochastic-based exponential random graph models (ERGMs) and stochastic actor-oriented models

(SAOMs). Although these models diverge in several ways, one of the key differences is the extent to which the interdependence of ties is considered.

Negative binomial gravity models are the most straightforward approach (Scherngell and Hu, 2011; Cao et al., 2018). Since the dependent variables – the number of inter-city collaborations – are usually count data with an over-dispersed distribution, a negative binomial regression is fitted. The gravity 'part' of the model allows studying the role of geography on the intensity of knowledge flows. A combination of both models can be applied to analyse individual choices or aggregated behaviour, i.e. to networks whose nodes are actors or cities (Scherngell, 2013). However, the model assumes ties to be independent which is not very realistic. Scherngell and Lata (2013) later extended the model with spatial filtering to consider the interdependence among ties sharing the same nodes.

MRQAP models are specifically designed to tackle tie interdependence by means of a combination of QAP and OLS/logistic regression (Li et al., 2021). The dependent and independent variables in MRQAP are in the form of matrices and the parameters are estimated by comparing regression statistics to the distribution of these statistics generated from row/column permutations of variables (Broekel et al., 2014). A recent example of its application can be found in the work of Park and Koo (2021), who use MRQAP to identify the impact of proximity on knowledge network formation in the Korean steel industry.

However, both above-mentioned models do not take endogenous structural variables into account, even though these may be of great importance (Liu et al., 2015; Zhang et al., 2016). To overcome this problem, ERGMs and SOAMs in social networks can be a powerful alternative to investigate how local structures influence network formation and this alongside other factors at the node and dyad levels. For instance, Qin et al. (2020) combine variables of local network structures, multidimensional proximity and node attributes in SOAMs, and describe the factors influencing China's knowledge collaboration networks. Broekel and Hartog (2013) confirm that the factors at the node, dyad, and structural network levels jointly determine the structure of intercity R&D collaboration network in Germany through ERGMs. However, previous model specifications were applied to binary networks. Unfortunately, scholars using SOAMs and traditional ERGMs have long been forced to 'dichotomize valued networks, resulting in significant loss of data. … The choice of where exactly to dichotomize a network is often based on arbitrary cutoffs, and can dramatically affect the results' (Pilny and Atouba, 2018). Recent advances in valued ERGMs address the important limitation (Krivitsky, 2012), which offers the potential to explore the mechanisms underlying the formation of intercity knowledge networks. However, to the best of our knowledge, this has not been explored in urban and regional science.

2.2 Determinants of intercity knowledge networks

Possible determinants of intercity knowledge networks can be categorized into exogenous and endogenous forces. The former includes forces at the node and the dyadic level, while the latter denotes factors at the network level.

Node-level factors can be city attributes including socio-economic development in the broadest sense of that term in general and the size of the knowledge base in particular. Cities with more innovative organizations and actors can be expected to have more linkages as they have more collaboration opportunities and demands. A similar argument can be applied to the socio-economic environment. Cities with better economic performance and a more open cultural and social environment would allow innovative agents to better overcome obstacles in the collaboration process (Yao et al., 2020). In addition, urban hierarchies in terms of administrative systems can also be a factor, especially in the Chinese socio-political context: cities with higher administrative levels can apply preferential policies to expand their resource bases, thus having the advantages over other cities for developing more intercity linkages (Cao et al., 2021a). Based on this, we propose the following hypothesis:

Hypothesis 1: The formation of intercity co-publication networks is influenced by city factors including economic performance, size of the knowledge base, and administrative level.

Dyad-level factors include edge attributes commonly discussed in the 'proximity literature' in economic geography, which has conceptual linkages with research in the homophily effect in sociology. The most straightforward of these is geographical proximity, which states that researchers in cities that are located closer to each other are more likely to collaborate due to easier face-to-face contact (Hoekman et al., 2009; Ma et al., 2014). Other forms of proximity are non-geographical, and are most commonly classified as social, institutional, cultural, and cognitive proximity. Social proximity can be defined as the relational embeddedness of researchers in terms of partnership, kinship, and friendship. Researchers already knowing each other more easily develop trust-based cross-regional collaborations, thus promoting knowledge exchange among cities (Breschi and Lissoni, 2009). Institutional proximity refers to the extent to which researchers or organizations operate with similar routines, established practices, and incentive structures (Boschma, 2005). Cultural proximity, in turn, broadly defined as informal institutional proximity, refers to the extent to which researchers share a coherent manner of interpretation and articulation such as a common language, religion, and ethnic community (Teixeira et al., 2008). Both forms of proximity can increase trust and lower transaction costs. As a corollary, institutional and cultural differences at the organization and city levels are often believed to be barriers to knowledge diffusion and collaborative activities (Ponds et al., 2007; Zhang et al., 2020b). In the field of scientific collaboration, cognitive proximity, sometimes sailing under the flag of technological proximity, refers to the degree of overlap between two researchers in terms of their knowledge bases and technological experience (Knoben and Oerlemans, 2006; Cao et al., 2019). It is argued that similarities in knowledge backgrounds of researchers can facilitate effective communication in knowledge collaboration and better absorption of external knowledge (Cao et al., 2021a). Hence, we posit the following hypothesis:

Hypothesis 2: The formation of intercity co-publication networks is influenced by geographical proximity as well as the main forms of non-geographical proximity.

And finally, structure-level factors are network configurations or local structures, which relate to properties of the entire network. Four factors are commonly examined in this context: edges, mutuality, k-star (preferential attachment), and triadic closure (transitivity) (cf. Ter Wal and Boschma, 2009; Broekel and Hartog, 2013; Zhang et al., 2020a; Qin et al., 2020). Edges are the most essential structural factor for networks, which is indispensable in stochastic-based models to help constraining the size of the networks in the simulation process. Mutuality may be a relevant variable in directed networks in that it reflects the probability of city pairs to develop a reciprocal relationship. However, given that we are dealing with undirected networks this is not a relevant indicator for our study. Preferential attachment in knowledge networks suggests that researchers or organizations who already have many linkages are more likely to attract or develop new collaborations in the future. When inter-personal and inter-organizational behaviours are aggregated at the intercity level, the overall network would show a hub-and-spoke structure. The preferential attachment processes have been shown in the case of intercity Internet networks (Vinciguerra et al., 2010) and firm networks (Li et al., 2021). Transitivity predicts that partners of researchers or organizations are more likely to engage in direct collaborations themselves, resulting in many triangles and dense cliques. Such cliques could also be seen as a sign of social proximity that could strengthen trust and willingness among researchers to conduct collaboration (Coleman, 2003). Thus, we put forward the following hypothesis:

> Hypothesis 3: The formation of intercity co-publication networks is influenced by endogenous effects such as preferential attachment and transitivity.

3 Research design

3.1 Construction of CIKNs

We use a comprehensive dataset of co-publications retrieved from the WoS to investigate the structure and dynamics of intercity knowledge collaboration in China. In line with Cao et al. (2021a, 2021b), CIKNs are aggregated at the scale of 217 prefectural-and-higher-level cities or municipalities across 20 city-regions including Hong Kong, Macau and Taiwan. In spite of the long-standing political conflicts between Taiwan and Mainland China, the inter-personal and inter-organizational exchanges in forms of scientific collaboration have witnessed a rapid increase in recent years, which leads us to include these collaborations in our data. The periods of 2002–2006 and 2012–2016 (2006 and 2016 for short) are selected because China experienced its most rapid annual increases in scientific publications during the period 2002–2006, ultimately replacing the United States as the largest provider of scientific publications in 2016 according to the National Center for Science and Engineering Statistics.[1] The average number of co-publications of each year in the study periods are used to smooth the fluctuations in an individual year. Consequently, two undirected and valued CIKNs are established for 2006 and 2016.

3.2 Characterization of intercity knowledge networks

3.2.1 Connectivity, degree, and density

Connectivity of a city is the sum of the city's co-publications with other cities, indicating the overall status of the city within the network. The degree of a city measures the number of partners/neighbours of the city in the network. Network density refers to the ratio between the actual number of edges and the total possible number of edges in the network. In an undirected network of size n, there will be $n(n-1)/2$ possible edges.

3.2.2 Average path length and clustering coefficient

The average path length (L) is the average number of steps along the shortest paths for all possible pairs of cities, while the average clustering coefficient (C) is the degree to which cities in the network tend to cluster together. They are specified as:

$$L = \frac{1}{n(n-1)} \sum\nolimits_{i,j=1}^{n} d_{ij}$$

$$C = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{2E_i}{k_i(k_i - 1)}$$

where n is the number of cities in the network, d_{ij} is the length of shortest path between city i and city j, k_i is the number of neighbours of city i, E_i is the number of edges between neighbours of city i.

3.2.3 Knowledge polycentricity

Knowledge polycentricity reflects the extent to which there are multiple knowledge centres as well as distributed linkages between them in the organization of intercity knowledge networks. We use a primacy-ratio-based indicator, whereby the knowledge polycentricity *(KP)* is defined as the ratio between the average connectivity of 2nd, 3rd, 4th, 5th, and 6th largest cities and the largest connectivity:

$$KP = \frac{1}{5} \sum\nolimits_{j=2}^{6} \mathrm{Con}_j / \mathrm{Con}_1$$

where Con_1 is the largest number of co-publications of cities and Con_j ($j =$ 2,3, … , 6) is the jth largest number of co-publications of cities.

3.2.4 Community detection

The CIKNs can be partitioned into different communities where cities are (more) densely connected internally but (more) sparsely connected externally. Modularity maximization is the most widely used method for community detection and it achieves this goal by iteratively optimizing local communities until the modularity is no longer improved (Blondel et al., 2008). The detecting technique can be implemented in R program by fast-greedy algorithm without an *a priori* threshold.

3.3 Exploring driving mechanisms: ERGM

3.3.1 An overview of ERGM

In order to reveal the mechanisms underlying CIKNs' formation and evolution, we employ an EGRM approach. ERGMs are stochastic models whose basic premise is that the global structure of the observed network is a snapshot of a set of ongoing and dynamic local processes (Lusher et al., 2013). An ERGM allows for flexible inclusion of such hypothesized processes and provides statistical inference on how these processes drive network formation by comparing the observed network to the theoretical sets. Since the CIKNs are weighted networks, valued ERGMs are more suitable for our simulation and can be expressed as:

$$\mathrm{P}_{\mathrm{r}\theta;h,\eta,g}(Y=y)=\frac{h(y)\exp\left(\eta(\theta)\cdot g(y)\right)}{k_{h,\eta,g}(\theta)}$$

where $\mathrm{P}_{\mathrm{r}\theta;h,\eta,g}(Y=y)$ is the probability that the network Y generated by the ERGM is identical to the observed network y, $g(y)$ is the hypothesized network statistic, $\eta(\theta)$ is the parameter corresponding to the network statistic, $k_{h,\eta,g}(\theta)$ is the normalized constant to ensure all probabilities summing up to 1, and $h(y)$ is the predetermined reference distribution which represents a baseline distribution of network linkages in the absence of any model terms.

Different from a simple Bernoulli distribution for binary ERGMs, the baseline distribution for valued ERGMs needs to take weights into consideration, thus determining the support and the basic shape of the ERGM distribution (Krivitsky, 2012). In practice, Poisson, geometric, binomial, and discrete uniform distributions are commonly used. The choice of reference distributions depends on the nature of data in the observed networks. Intercity co-publications are count data exhibiting over-dispersion (i.e. they have a variance larger than the mean), and therefore a binomial distribution is specified as the reference distribution for modelling CIKNs, which is also demonstrated to be typical for measuring strength, frequency, and intensity (Pilny and Atouba, 2018).

The estimation of the parameters can be implemented through a Markov Chain Monte Carlo (MCMC) algorithm as this is often the most accurate approach (van Duijn et al., 2009). Starting with stochastic simulation from a set of initial parameters, MCMC iteratively refines those values by comparing the simulated networks with the observed one until the parameters generate networks bearing high resemblance with the observed network and the parameter estimates stabilize. Poorly specified ERGMs fail to converge and can be fixed by removing factors that are problematic.

3.3.2 Specification of variables

The dependent variable in our ERGMs is the observed CIKN at one time point, while the independent variables incorporate possible exogenous node-level and dyad-level factors and endogenous network-level factors as reviewed in section 2.2. To test hypothesis 1, we employ the gross domestic product *(GDP*), and *R&D* investment from local government, the number of *universities*, the number of college *students* as the proxy accounting for economic performance and knowledge base, respectively. These data have been drawn from the China City Statistical Yearbook published by the National Bureau of Statistics for mainland cities and from the official statistics websites for Taiwanese cities. The different variables represent the mean values of a five-year moving window of 2002–2006 and 2012–2016. In addition, Provincial *capital* is a dummy variable to measure the impact of administrative hierarchy.

Building on the early ideas presented by the French School of proximity dynamics (cf. Rallet and Torre, 1999; Kirat and Lung, 1999), Boschma's (2005) summary of multi-dimensional proximity has become particularly popular in economic geography. It entails geographical, institutional, organizational, social and cognitive proximity. These well-known forms of proximity are nevertheless not always evident in intercity co-publication networks and some of them are intertwined in our empirical case. For instance, when interpersonal scientific collaboration between scholars from different organizations such as universities, academic institutes, companies, and hospitals is aggregated into the intercity collaboration, the original organization (affiliation) information would be disregarded. Instead, the 'affiliation' of a city would be the province or region where it is situated, in which sense the organizational proximity of two cities can also be taken as institutional proximity (Broekel, 2015) because cities in the same province or city region are often subject to the same institutional framework at the macro level. Alongside institutional and organizational proximity, social proximity can also play a key role in the scientific collaboration among cities. However, it shows strong overlap with the key endogenous force of triadic closure as Ter Wal (2014) demonstrated that triadic closure is a specific case of the social proximity effect on tie formation. In addition, Broekel (2015) argued that social and cognitive proximity between linked actors are quite likely to be correlated. To avoid possible multicollinearity and model degeneracy, we restrict our treatment of non-geographical forms of proximity to institutional proximity and cultural proximity.

As a result, to test hypothesis 2, our proximity framework includes geographical, institutional, and cultural dimensions. Geographical proximity is measured by the Euclidean *distance* between cities. Institutional proximity is measured by two variables of *province* and *region*, which are dyadic dummy variables with cities in the same province or city-region being 1, and 0 otherwise. The scope of city-regions is based on the urban agglomerations in the 13^{th} fifth plan of China as well as Cao et al. (2021a). Cultural proximity is also dyadic dummy variable in which city-pairs sharing dialect are 1 and 0 otherwise based on 2010 Atlas of Chinese Dialects (Xiong and Zhang, 2012). To test hypothesis 3, variables of edges, k-star, triadic closure are considered except mutuality because the CIKNs are undirected. In a valued ERGM, the three variables correspond to *sum*, *nodesqrtcovar*, and *transitiveweights*.

According to the specification of models and variables, we run the ERGM simulation through the packages of statnet and ergm.count in the R program. The estimates of parameters are tested by p-value and an overall goodness-of-fit (GOF) is given by the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The best fitting model can be achieved by performing various types of model comparison, of which a smaller AIC and BIC indicate a better GOF.

4 Results and discussion

4.1 The evolving structures of CIKNs

Figure 1 shows the main structure of CIKNs in 2006 and 2016, in which the size of nodes denotes the total number of co-publications of cities, the width of edges denotes the collaboration intensity of city-pairs, and the colour of nodes denotes the communities each city belongs to. With regard to topological properties, it is clear that the scale and scope of CIKNs have dramatically increased from 2006 to 2016. Compared with the CIKN in 2006 including 2688 intercity collaboration linkages among 192 cities, the network has intensified with 217 cities engaging in 7820 knowledge collaboration process in 2016. Therefore, the network density has risen from 0.147 to 0.337, with average degree growing from 28 to 72. In 2006, it was still quite difficult to initiate knowledge collaboration among cities because the average path length was 1.920, which was greater than that

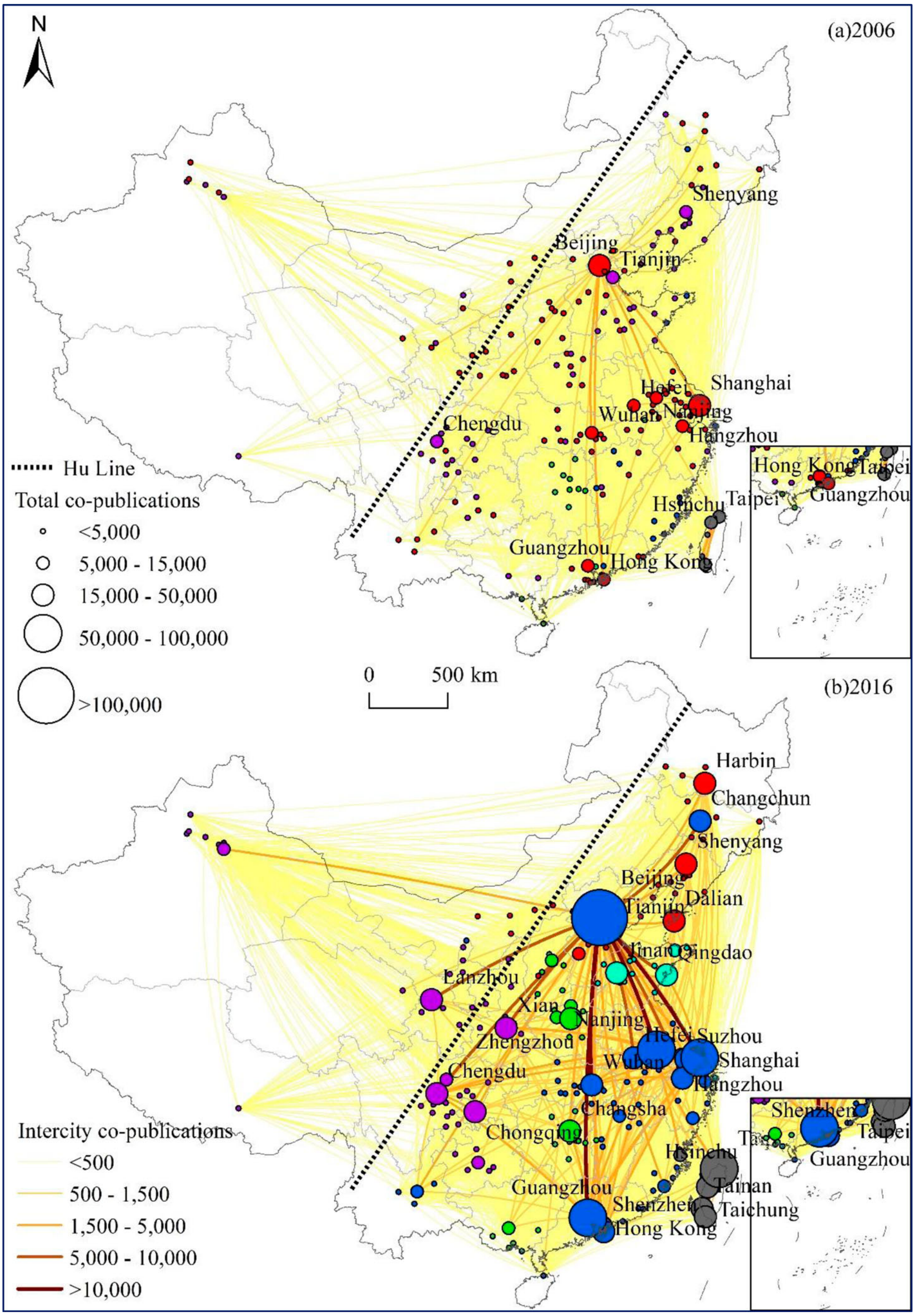

Figure 1. China's intercity knowledge networks, 2006–2016.

of a random network of the same size. However, the length was shortened to 1.576 by 2016, smaller than that of an identical-size random network. Meanwhile, the average clustering coefficient ascended from 0.449 to 0.598. The smaller average path length together with the larger clustering coefficient indicate a small-world structure of CIKN in 2016, which can be deemed beneficial for innovation output because the inter-organizational and inter-personal trust will increase within a more closed and embedded social contexts (Uzzi and Spiro, 2005; Fleming et al., 2007).

With regard to the overall collaboration landscape, intercity co-publications in both 2006 and 2016 were dense in the southeastern but more dispersed in the northwestern parts of China. This broad division is most clearly represented by the so-called Hu Line, an imaginary line stretching from Heihe (a northern city located on the Russian border) to Tengchong (a southwestern city bordering Myanmar) (Chen et al., 2019). This patten has also been found in China's intercity transport networks (Wang et al., 2011a), population networks (Zhang et al., 2020a), corporate networks (Pan et al., 2017), and to a large extent reflects the uneven development of China's space-economy. More specifically, the overall CIKN in 2006 was mainly characterized as a hub-and-spoke structure centred on Beijing and the Beijing-Shanghai/Nanjing-Hong Kong/Guangzhou triangles, alongside dense connections among cities in Taiwan. In 2016, Beijing's knowledge flows radiating outward was strengthened and the knowledge spillover to cities in the northeast and northwest regions became apparent. Apart from that, the backbone of the CIKN in 2016 has transferred to a diamond-shaped structure anchored by four country-level urban agglomerations, i.e. the Beijing–Tianjin–Hebei, the Yangtze River Delta, the Pearl River Delta, and the Chengdu–Chongqing urban agglomerations, which is also observed in China's intercity patent cooperation network (Duan et al., 2018). Consequently, the degree of knowledge polycentricity has increased from 0.298 in 2006 to 0.331 in 2016. Alongside this broader pattern, the influence of the top-down administrative system is obvious in the spatial configuration of the CIKNs, especially in the Mainland, which is in line with Andersson et al.'s (2014) findings. As can be seen in Table 1, apart from Taiwanese cities, the top 20 cities in terms of co-publications were all provincial capitals and special administrative regions in 2006 and were provincial capitals, special administrative regions, and one sub-provincial city (Shenzhen) in 2016.

Table 1. Top-20 cities of co-publications in China's intercity knowledge networks.

Rank	City	Co-publications in 2006	City	Co-publications in 2016
1	Beijing	34,122	Beijing	194,096
2	Shanghai	15,293	Shanghai	91,124
3	Taipei	10,773	Nanjing	67,894
4	Nanjing	9583	Guangzhou	62,574
5	Hong Kong	7658	Taipei	51,212
6	Wuhan	7595	Wuhan	48,002
7	Guangzhou	6117	Hangzhou	41,517
8	Hangzhou	6061	Chengdu	38,811
9	Hsinchu	5350	Xi'an	35,277
10	Tianjin	5249	Tianjin	31,973
11	Kaohsiung	5200	Hefei	30,872
12	Hefei	5166	Hong Kong	28,991
13	Chengdu	5131	Ji'nan	28,422
14	Shenyang	5074	Changsha	28,152
15	Taichung	4760	Changchun	24,525
16	Xi'an	4692	Shenzhen	23,929
17	Tainan	4638	Shenyang	22,365
18	Changchun	4157	Chongqing	21,720
19	Changsha	4002	Harbin	21,263
20	Lanzhou	3896	Taichung	21,258

With regard to the hierarchical tendencies, although Beijing and Shanghai were consistently leading knowledge collaboration centres in China, subsequent ranks changed a lot (Table 1). The connectivity of cities in mainland China has surged in the CIKN. First, Hong Kong's leading role in the Guangdong-Hong Kong-Macao Greater Bay Area (GBA) was replaced by Guangzhou in 2016. In addition, Shenzhen entered into the Top-20 in 2016 and emerged as another regional centre in the GBA science system, which can be attributed to the establishment of cross-city branches of many renowned universities and research institutes such as Peking University, Harbin Institute of Technology and Sun Yat-sen University since the early 2000s (Ma et al., 2021). Second, notwithstanding Taiwanese cities such as Taipei, Hsinchu, and Taichung producing more co-publications in 2016 than 2006, the growth is much smaller than in mainland cities, leading to a drop in the ranking. Third, the position of Nanjing, Wuhan, Hangzhou, Tianjin was relatively stable whereas cities in the western region like Chengdu, Xi'an, Chongqing and those in the eastern region like Changchun, Shenyang, Harbin gained much prominence in 2016. This is greatly associated with the strategies of development of west China, revitalization of northeast China, and rise of central China.

Community detection divides cities into five coherent subgroups in 2006 and six in 2016. There exists some stability as well as changes in these patterns. Although the east-strait (Taiwanese) community is relatively stable, an inclusion of Fuzhou, the capital of Fujian province in 2016 suggests deepening cross-strait interactions as well as a possible effect of geographical proximity. The formation of Shandong Peninsula community in 2016 shows a 'provincial bias' (Andersson et al., 2014; Li and Phelps, 2017), that is, cities in Shandong Province are more likely to develop scientific collaborations at the intra-provincial level than at the inter-provincial level. Such spatial patterns are also found at the city-region level. For instance, cities in the northeast part that were part of the Beijing community in 2006 formed a new community in 2016, which coincides with the development of Harbin-Changchun and Central South Liaoning urban agglomerations. Jointly with the spatial political bias, the impact of geographical and cultural proximities can to some extent be observed in the combination of cities in central Shanxi, central Plain, and the middle reach of Yangtze River regions into a central community, and the expansion of the Chengdu-Chongqing community by including nearby cities in 2016.

4.2 The driving mechanisms behind CIKNs

Table 2 summaries the results of the ERGMs, including all specified variables for CIKNs in 2006 and 2016 and which are also the best fitting models. The models are stable and converge. A brief look at significant variables in both periods suggests the three proposed hypotheses all hold in 2016 even though the knowledge base characterized by the number of students and the cultural proximity characterized by dialect were insignificant in 2006.

Regarding Hypothesis 1, as expected, cities with better economic performance, larger R&D investment, more universities, higher administrative level tend to have more knowledge collaboration: the coefficients for *GDP*, *R&D*, *universities*, and *capital* are all positive and significant at the 1% level. Nevertheless, the number of *students* in a city does not impact the structure of intercity knowledge exchange in 2006 whereas the situation has changed in 2016, suggesting that students play an increasingly pivotal

Table 2. Results of ERGMs for CIKNs in 2006 and 2016.

	Model 2006				Model 2016			
Varibles	Estimate	Std. error	*p*-value	Sign.	Estimate	Std. error	*p*-value	Sign.
Exogenous forces								
(1) Node level								
GDP	0.0007	0.0001	<1e-04	***	0.0013	0.0000	<1e-04	***
R&D	0.1735	0.0246	<1e-04	***	1.0600	0.4701	<1e-04	***
Universities	0.0659	0.0061	<1e-04	***	0.0347	0.0026	<1e-04	***
Students	0.0023	0.0044	0.6105		0.0185	0.0029	<1e-04	***
Capital	1.3270	0.0855	<1e-04	***	1.2200	0.0074	<1e-04	***
(2) Dyad level								
Distance	−0.0003	0.0000	<1e-04	***	−0.0002	0.0000	<1e-04	***
Province	1.6280	0.0440	<1e-04	***	0.8708	0.0120	<1e-04	***
Region	0.2394	0.0465	<1e-04	***	0.4084	0.0101	<1e-04	***
Dialect	0.0296	0.0647	0.6469		0.0935	0.0042	<1e-04	***
Endogenous forces								
(3) Structural network level								
Sum	−5.0010	0.0466	<1e-04	***	−4.4660	0.0498	<1e-04	***
Nodesqrtcovar	0.4421	0.0564	<1e-04	***	0.3200	0.0144	<1e-04	***
Transitiveweights	0.1047	0.0541	0.0528	*	0.5570	0.0495	<1e-04	***
Akaike Inf. Crit. (AIC)		−59682				−46718		
Bayesian Inf. Crit. (BIC)		−59591				−46624		

***0.01, **0.05,*0.1.

role in scientific collaboration. The coefficients of *GDP* show an upward trend, but the values are still fairly small, which confirms Wang et al.'s (2021) work for the collaborative innovation network of the Yangtze River Delta region. In contrast, the coefficients for *universities* have decreased, albeit that the impact of universities on promoting knowledge collaboration is still bigger than students and GDP. This might be because universities are more suitable to represent the main and direct producers of co-publications while most students are not capable to conduct scientific collaboration on the one hand and, on the other hand, GDP is a more general indicator covering many aspects beyond scientific research. Provincial capital effects on intercity knowledge networks are strongest, followed by R&D investment from local government which has risen sharply from 2006 to 2016. This implies that knowledge collaboration and innovation activities are, to a large extent, influenced by government policies (Wang et al., 2011). For example, the scientific fields in which the government invests more, or encourages in its official documents, will undoubtedly attract more research interest and thereby might stimulate more collaboration. A slight drop in *capital*'s coefficients from 2006 to 2016 indicates a decentralization trend in knowledge flows, confirming the previous descriptions of structural evolution of CIKNs.

Regarding Hypothesis 2, geographical proximity has the least influence on promoting the intercity scientific collaboration and the impact of distance has weakened from 2006 to 2016. This is consistent with recent studies on geography and proximity in intercity flows. That is, although still exerting an influence, the distance barriers become less and less marked due to the development of communication technology and improvement of transport infrastructures in China (Zhang et al., 2020b). The effects of institutional proximity are most salient. Although collaborative activities can physically take place across regions, they still tend to follow the territorial configuration of political space (cf. Ma, 2005) as this instills an identical framework of collaboration policies, fund allocation, and output management. The coefficients of *province* are bigger than that of *region*, suggesting that cities are more likely to collaborate with other cities in the same province than those in the same city-region. However, with the increase of *region*'s coefficients and the decrease of *province*'s coefficients, the effects of two institutional proximities are converging and the integration of urban agglomerations in China is enhanced. Cultural proximity is an exception since it does not impact the structure of intercity knowledge exchange in 2006 but does so in 2016. The instability of the cultural proximity effect is also found in Yang et al. (2018). The significant and positive effect in 2016 can be associated with the development of China's urban agglomerations.

Regarding Hypothesis 3, variables' coefficients are significant in both models, so the effects of preferential attachment and transitivity are robust, indicating the need to include endogenous forces in the network modelling. Well-connected cities are more prone to have additional collaboration than sparsely connected cities. This endogenous effect is particularly evident in the reinforced hub-and-spoke structure centred on Beijing and the rise of regional knowledge centres. Primary cities of each province prefer developing scientific collaboration with Beijing rather than with cities in their own province. Following the same logic, other cities have more co-publications with their provincial capitals than small nearby cities. This mechanism would help cities to gain more knowledge spillovers at smaller costs. The decrease of coefficients signals that the preferential attachment process has weakened from 2006 to 2016, which parallels the aforementioned decline in capital effects and the polycentric development of knowledge centres in China. Meanwhile, we find a growing, significant, and positive impact of triadic closure, which corresponds to the visual inspection of CIKNs in the Figure 1. For example, with the intensifying collaboration of Beijing-Shanghai and Beijing-Guangzhou, the open triad has evolved into a closed one from 2006 to 2016 despite not sharing geographical, institutional, and cultural proximity. In this regard, the structural force played a crucial role in the formation of diamond-shaped backbone of Beijing-Shanghai-Guangzhou-Chengdu. Broekel and Hartog (2013) and Qin et al. (2020) also found that triadic closure is an important driving force in the network formation of German's R&D collaboration network and China's knowledge collaboration network, respectively. The growth of coefficients indicates actors in cities increasingly utilize network resources by forming collaborative links with partners of their partners. The tendency towards triadic closure yields dense cliques of strongly interconnected cities and the social cohesion in the strong cliques in turn fosters the trust and lessens the costs in knowledge exchange (Ter Wal, 2014). The coefficient of *sum* is negative and significant, which is a common feature of networks and means that the observed network tends to be less dense than exponential random networks.

5 Conclusions

In this research, we examined the topological and spatial structures as well as dynamics of CIKNs based on co-publications from WoS for the periods 2002–2006 and 2012–2016. We then explored the driving mechanisms of the CIKN's formation and evolution using an exponential random graph modelling approach. Our main conclusions are as follows:

First, the intercity scientific collaborations have dramatically increased from 2006 to 2016, resulting in a much denser network with a shorter average path length and a larger clustering coefficient. The overall geography is relatively stable in that co-publications are denser in the southeastern and sparser in the northwest part of China. Beijing remains the dominant knowledge centre, but with the emergence of regional centres in mainland China, the diamond-shaped structure anchored by four country-level urban agglomerations comes to the fore in 2016, showing an increasingly polycentric development and leading to the formation of more local communities in the CIKN.

Second, the ERGM results suggest that preferential attachment and transitivity are significant endogenous processes driving the formation of the CIKNs. As for exogenous factors, administrative level and R&D play the strongest and significant role in intercity scientific collaborations. GDP's impact is limited and can almost be ignored, and so does geographical proximity. Institutional proximity is the most important among the different proximity effects. Provincial preference in intercity scientific collaborations is still stronger than city-regional preference. The endogenous and exogenous forces can to some extent substitute for one another.

Third, the comparison of coefficients in 2006 and 2016 indicates that the preferential attachment effect has weakened whereas the transitivity effect has strengthened. The evolution of the CIKNs is to some degree a trade-off between these effects. There is a slight drop of administrative effect but a surge in R&D's influence. Students and cultural proximity play an increasing pivotal role in scientific collaboration. With the rise of *region*'s coefficients and the decline of *province*'s coefficients, the effects of two institutional proximities are converging, suggesting an increasing integration of urban agglomerations in China.

Our research contributes to the literature by simultaneously considering exogenous forces and endogenous forces, and extending the application of ERGMs from binary networks to valued networks. The exogenous forces underpinning the intercity scientific collaboration have been shown in the extensive literature on knowledge spillovers and urban innovation systems. The endogenous forces unveiled in this study are more innovative. At the early stage of scientific collaboration, cities tend to develop relationship with well-connected cities, which allows obtaining knowledge spillover efficiently. The preferential attachment process produces many open triads. When the collaboration in an open triad is strengthened to the extent that it forms a high-trust environment, the triadic closure works and the interconnected cliques come to the fore.

There are obviously also a number of limitations to our study, which might open perspectives for future research. First, given the inclusion of Hong Kong, Macau and Taiwan, we employed the co-publication data retrieved from the WoS rather than the Chinese domestic databases. Publishing in WoS indexed journals is on average more difficult than publishing in 'domestic' journals, and authors publishing in English thus tend to somewhat affiliated to higher-level universities located in larger cities. A careful examination of scientific collaboration of cities in mainland China calls for an incorporation and comparison of different datasets in both Chinese and English. Second, different from previous (spatial) econometric models that can detect multicollinearity and solve endogeneity effectively, ERGMs do not deal with these problems before simulation. An alternative is to specify many versions of ERGMs by stepwise-adding variables and then to check whether the models converge. When a newly added variable causes model degeneracy, the variable can be removed. However, due to the computational complexity, it is neither judicious nor possible to include as many candidate determinants as possible into the model because the simulation for large networks is quite time-consuming when the triadic closure and preferential attachment variables are incorporated. Even though the retained variables lead to model convergence, we cannot be sure if there is multicollinearity among them. In this sense, the absolute values of estimated parameters for each factor are not that useful, but comparisons of parameters for different factors and for the same factors in different time periods can provide insights into the importance of factors that drive the network formation and evolution.

Note

1. Data source: https://ncses.nsf.gov/pubs/nsb20206/publication-output-by-region-country-or-economy

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work is supported by the National Natural Science Foundation of China (41901189, 52008298), Natural Science Foundation of Jiangsu Province (BK20190797), China Postdoctoral Science Foundation (2020M671229), Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (19KJB170016) and KU Leuven C1 Project (C14/21/021).

ORCID

Liang Dai http://orcid.org/0000-0001-7421-8203
Ben Derudder http://orcid.org/0000-0001-6195-8544
Zhan Cao http://orcid.org/0000-0002-6439-3013

References

Andersson, D. E., Gunessee, S., Matthiessen, C. W., & Find, S. (2014). The geography of Chinese science. *Environment and Planning A*, *46*(12), 2950–2971.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of community hierarchies in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *10*, P10008.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61–74.

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, *9*(4), 439–468.

Broekel, T. (2015). The co-evolution of proximities – a network level study. *Regional Studies*, *49*(6), 921–935.

Broekel, T., Balland, P. A., Burger, M., & Oort, F. V. (2014). Modeling knowledge networks in economic geography: A discussion of four methods. *Annals of Regional Science*, *53*(2), 423–452.

Broekel, T., & Hartog, M. (2013). Determinants of cross-regional R&D collaboration networks: An application of exponential random graph models. In T. Scherngell (Ed.), *The geography of networks and R&D collaborations* (pp. 49–70). Berlin, NY: Springer.
Cao, Z., Derudder, B., Dai, L., & Peng, Z. (2021a). 'Buzz-and-pipeline' dynamics in Chinese science: The impact of interurban collaboration linkages on cities' innovation capacity. *Regional Studies*, (6), 1–17. doi:10.1080/00343404.2021.1906410
Cao, Z., Derudder, B., & Peng, Z. (2018). Comparing the physical, functional and knowledge integration of the Yangtze River Delta city-region through the lens of inter-city networks. *Cities*, *82*, 119–126.
Cao, Z., Derudder, B., & Peng, Z. (2019). Interaction between different forms of proximity in inter-organizational scientific collaboration: The case of medical sciences research network in the Yangtze River Delta region. *Papers in Regional Science*, *98*(5), 1903–1924.
Cao, Z., Peng, Z., & Derudder, B. (2021b). Interurban scientific collaboration networks across Chinese city-regions. *Environment and Planning A: Economy and Space*, *53*(1), 6–8.
Chen, D., Zhang, Y., Yao, Y., Hong, Y., Guan, Q., & Tu, W. (2019). Exploring the spatial differentiation of urbanization on two sides of the Hu Huanyong Line–based on nighttime light data and cellular automata. *Applied Geography*, *112*, 102081.
Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Cambridge, MA: Harvard Business School Press.
Coleman, J. S. (2003). *Social capital in the creation of human capital*. In R. Cross, A. Parker, & L. Sasson (Eds.), *Networks in the knowledge economy* (pp. 57–81). Oxford: Oxford University Press.
Duan, D., Du, D., Chen, Y., & Zhai, Q. (2018). Spatial-temporal complexity and growth mechanism of city innovation network in China. *Scientia Geographica Sinica*, *38*(11), 1759–1768.
Fleming, L., King C. III, & Juda, A. I. (2007). Small worlds and regional innovation. *Organization Science*, *18*(6), 938–954.
Florida, R., Adler, P., & Mellander, C. (2017). The city as innovation machine. *Regional Studies*, *51* (1), 86–96.
Gui, Q., Liu, C., & Du, D. (2019). Globalization of science and international scientific collaboration: A network perspective. *Geoforum*, *105*, 1–2.
Hoekman, J., Frenken, K., & Van Oort, F. (2009). The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, *43*(3), 721–738.
Kirat, T., & Lung, Y. (1999). Innovation and proximity: Territories as loci of collective learning processes. *European Urban and Regional Studies*, *6*(1), 27–38.
Knoben, J., & Oerlemans, L. A. (2006). Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, *8*(2), 71–89.
Krivitsky, P. N. (2012). Exponential-family random graph models for valued networks. *Electronic Journal of Statistics*, *6*, 1100.
Lata, R., Scherngell, T., & Brenner, T. (2015). Integration processes in European research and development: A comparative spatial interaction approach using project based research and development networks, co-patent networks and co-publication networks. *Geographical Analysis*, *47*(4), 349–375.
Li, D., Wang, T., Wei, Y. D., & Yuan, F. (2015). Spatial and temporal complexity of scientific knowledge network and technological knowledge network on China's urban scale. *Geographical Research*, *34*(3), 525–540.
Li, L., Derudder, B., Shen, W., & Kong, X. (2021). Exploring the dynamics of the disaggregated intercity corporate network in the Yangtze River Delta, China: A relational event approach. *Journal of Geographical Systems*, 1–26. https://doi.org/10.1007/s10109-021-00358-2.
Li, Y., & Phelps, N. (2017). Knowledge polycentricity and the evolving Yangtze River Delta megalopolis. *Regional Studies*, *51*(7), 1035–1047.
Liefner, I., & Hennemann, S. (2011). Structural holes and new dimensions of distance: The spatial configuration of the scientific knowledge network of China's optical technology sector. *Environment and Planning A*, *43*(4), 810–829.

Liu, X., Derudder, B., & Liu, Y. (2015). Regional geographies of intercity corporate networks: The use of exponential random graph models to assess regional network-formation. *Papers in Regional Science*, *94*(1), 109–126.
Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential random graph models for social networks: Theory, methods, and applications*. Cambridge: Cambridge University Press.
Ma, H., Fang, C., Pang, B., & Li, G. (2014). The effect of geographical proximity on scientific cooperation among Chinese cities from 1990 to 2010. *PloS One*, *9*(11), e111705.
Ma, H., Li, Y., & Huang, X. (2020). Proximity and the evolving knowledge polycentricity of megalopolitan science: Evidence from China's Guangdong-Hong Kong-Macao Greater Bay area, 1990–2016. *Urban Studies*, *58*(12): 2405–2423.
Ma, L. J. (2005). Urban administrative restructuring, changing scale relations and local economic development in China. *Political Geography*, *24*(4), 477–497.
Matthiessen, C. W., Schwarz, A. W., & Find, S. (2010). World cities of scientific knowledge: Systems, networks and potential dynamics. An analysis based on bibliometric indicators. *Urban Studies*, *47*(9), 1879–1897.
Pan, F., Bi, W., Lenzer, J., & Zhao, S. (2017). Mapping urban networks through inter-firm service relationships: The case of China. *Urban Studies*, *54*(16), 3639–3654.
Park, S., & Koo, Y. (2021). Impact of proximity on knowledge network formation: The case of the Korean steel industry. *Area Development and Policy*, *6*(2), 181–199.
Pilny, A., & Atouba, Y. (2018). Modeling valued organizational communication networks using exponential random graph models. *Management Communication Quarterly*, *32*(2), 250–264.
Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, *86*(3), 423–443.
Qin, L., Teng, T., Zhang, Y., & Zeng, G. (2020). Evolution's characteristics and influence factors of China's university knowledge collaboration network. *Science & Technology Progress and Policy*, *37*(22), 125–133.
Rallet, A., & Torre, A. (1999). Is geographical proximity necessary in the innovation networks in the era of global economy? *GeoJournal*, *49*(4), 373–380.
Scherngell, T. (2013). *The geography of networks and R&D collaborations*. Berlin, NY: Springer.
Scherngell, T., & Hu, Y. (2011). Collaborative knowledge production in China: Regional evidence from a gravity model approach. *Regional Studies*, *46*(6), 755–772.
Scherngell, T., & Lata, R. (2013). Towards an integrated European research area? Findings from eigenvector spatially filtered spatial interaction models using European framework programme data. *Papers in Regional Science*, *92*(3), 555–577.
Snijders, T. A., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology*, *36*(1), 99–153.
Snijders, T. A., Van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, *32*(1), 44–60.
Teixeira, A. A., Santos, P., & Oliveira Brochado, A. (2008). International R&D cooperation between low-tech SMEs: The role of cultural and geographical proximity. *European Planning Studies*, *16*(6), 785–810.
Ter Wal, A. L. (2014). The dynamics of the inventor network in German biotechnology: Geographic proximity versus triadic closure. *Journal of Economic Geography*, *14*(3), 589–620.
Ter Wal, A. L., & Boschma, R. A. (2009). Applying social network analysis in economic geography: Framing some key analytic issues. *The Annals of Regional Science*, *43*(3), 739–756.
Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, *111*(2), 447–504.
van der Wouden, F., & Rigby, D. L. (2019). Co-inventor networks and knowledge production in specialized and diversified cities. *Papers in Regional Science*, *98*(4), 1833–1853.
van Duijn, M. A., Gile, K. J., & Handcock, M. S. (2009). A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models. *Social Networks*, *31*(1), 52–62.

Vertes, P. E., Alexander-Bloch, A. F., Gogtay, N., Giedd, J. N., Rapoport, J. L., & Bullmore, E. T. (2012). Simple models of human brain functional networks. *Proceedings of the National Academy of Sciences*, *109*(15), 5868–5873.

Vinciguerra, S., Frenken, K., & Valente, M. (2010). The geography of internet infrastructure: An evolutionary simulation approach based on preferential attachment. *Urban Studies*, *47*(9), 1969–1984.

Wang, H., Sun, Q., Guo, J., & Du, M. (2021). Research on evolution dynamics of collaborative innovation network in the urban agglomeration of Yangtze River Delta based on ERGM. *Science & Technology Progress and Policy*, *38*(14), 45–53.

Wang, J., Mo, H., Wang, F., & Jin, F. (2011a). Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. *Journal of Transport Geography*, *19*(4), 712–721.

Wang, T., Hennemann, S., Liefner, I., & Li, D. (2011b). Spatial structure evolution of knowledge network and its impact on the NIS: Case study of biotechnology in China. *Geographical Research*, *30*(10), 1861–1872.

Xiong, Z., & Zhang, Z. (2012). *Language Atlas of China* (2nd ed.). Beijing: The Commercial Press.

Yang, W., Du, D., Ma, Y., & Jiao, M. (2018). Network structure and proximity of the trade network in the belt and road region. *Geographical Research*, *37*(11), 2218–2235.

Yao, L., Li, J., & Li, J. (2020). Urban innovation and intercity patent collaboration: A network analysis of China's national innovation system. *Technological Forecasting and Social Change*, *160*, 120185.

Zhang, S., Derudder, B., & Witlox, F. (2016). Dynamics in the European air transport network, 2003–9: An explanatory framework drawing on stochastic actor-based modeling. *Networks and Spatial Economics*, *16*(2), 643–663.

Zhang, W., Chong, Z., Li, X., & Nie, G. (2020a). Spatial patterns and determinant factors of population flow networks in China: Analysis on Tencent location big data. *Cities*, *99*, 102640.

Zhang, W., Derudder, B., Wang, J., & Witlox, F. (2020b). An analysis of the determinants of the multiplex urban networks in the Yangtze River delta. *Tijdschrift voor Economische en Sociale Geografie*, *111*(2), 117–133.