INTRODUCTION

Since the publication of the seminal work of Griliches (1979), knowledge production has been an important topic in economic geography and regional science (Audretsch & Keilbach, 2004; Charlot, Crescenzi, & Musolesi, 2015; Crescenzi & Jaax, 2017; Jaffe, 1989; Lee, 2017; Miguelez & Moreno,
2013; Ó hUallacháin, & Leslie, 2007; Ponds, Oort, & Frenken, 2010). The knowledge production model assumes that the functional relationship between knowledge output and knowledge input is linear, the former mainly depends on a set of inputs, such as research and development expenditure (RDE) or human capital (Charlot et al., 2015; Lee, 2017). The new growth theory and knowledge-based economy have emphasized the importance of knowledge production or science and technology, which are regarded as the sources of economic growth (Lucas, 1988; Nelson, 2000; OECD, 1996; Romer, 1986). Therefore, knowledge and innovation are increasingly becoming the key for firms, cities, regions, and countries to maintain competitive advantages and sustainable development. However, knowledge does not stop flowing as spatial distances and territorial borders (Audretsch & Feldman, 2004; Pan, Kaski, & Fortunato, 2012). There is consensus among scholars that knowledge spillovers occur not only on the local or regional level but also on the national or even international scale (Bathelt, Malmberg, & Maskell, 2004; Gui, Liu, & Du, 2018; Hoekman, Frenken, & Oort, 2009; McKelvey, Alm, & Riccaboni, 2003; Scherngell & Hu, 2011).

When it comes to knowledge spillovers, there are two different research directions. One stream of the literature has adopted the regional knowledge production function approach based on the spatial econometrics techniques, which has been taken as a powerful tool for capturing spatial spillover effects (Basile & Minguez, 2018) and is based on the assumption that geography is a channel for spillovers (Charlot et al., 2015; Ó hUallacháin, & Leslie, 2007; Ponds et al., 2010). The related literature has initially highlighted the spatial weight matrix, the parametric estimation methods and the importance of tangible inputs, however, recently the focus has shifted to network weight matrix, the semi-parametric estimation methods, and intangible factors (Basile & Minguez, 2018; Charlot et al., 2015; Hazir, Lesage, & Autant-Bernard, 2016; Lee, 2017; Maggioni et al., 2007; Miguelez & Moreno, 2013; Ponds et al., 2010). Another stream has taken social network analysis technique, a promising tool for capturing the structure and dynamics of relational spillovers effects, to enrich the literature on knowledge spillovers (Araújo, Gonçalves, & Taveira, 2018; Boschma & Ter Wal, 2007; Breschi & Lenzi, 2015, 2016; Fleming, King, & Juda, 2007; Gluckler, 2007; Ter Wal & Boschma, 2009). The related literature has focused on the influence of network properties on innovation performance or inventive productivity, such as external linkages and gatekeepers (Breschi & Lenzi, 2015), internal reach, external reach and clique density (Araújo et al., 2018; Breschi & Lenzi, 2016), network density (De Noni, Orsi, & Belussi, 2018), betweenness centrality (Broekel, Bracht, Duschl, & Brenner, 2017), or small-world structure (Fleming et al., 2007; He & Fallah, 2014). As Boschma (2005) put it, geographical proximity is neither necessary nor sufficient conditions for knowledge spillovers to occur, while network relations may act as transmission channels for knowledge diffusion (Bathelt et al., 2004; Breschi & Lissoni, 2009; Maggioni et al., 2007; Miguelez & Moreno, 2013; Ponds et al., 2010). There is a growing body of literature recognizing the significance of network structure (Araújo et al., 2018; Crespo, Suire, & Vicente, 2014; De Noni et al., 2018; He & Fallah, 2014). Following this logic, taking only the spatial spillover effects into consideration and neglecting the role of network structure are insufficient.

In line with the second stream of literature, the objective of this article is to analyze the influence of network structure on knowledge production in collaboration network. With the globalization of science as the research background and a country as the unit of analysis, this paper constructs international scientific collaboration network based on publication coauthorship data from the Clarivate Analytics’ Web of Science database in the period from 2000 to 2015. By adopting knowledge production function framework and social network analysis, we investigate the effect of three structural features in the collaboration network—degree centrality, structural holes, and small-world quotient—on national knowledge production performance. Our findings show that the coefficients of the three
network properties are positive and statistically significant in panel data estimates, which means that higher degree centrality, structural holes, and small-world quotient will produce better future knowledge output.

The main contribution of this paper is threefold. First, economic geography and regional science has emphasized the importance of network structure or network position, while there is a scarcity of empirical analysis about the relationship between social networks and knowledge production performance. This study could substantially further our understanding of knowledge production. Second, this article unites both of the two perspectives: knowledge production function and social network analysis, which emphasizes the complementarity between internal inputs and external resources. Third, copublications reflect knowledge flows better than copatents (Li & Phelps, 2017). Technological collaboration networks using copatent data have been extensively discussed in previous studies, while scientific collaboration networks using copublication data have received only limited attention.

The rest of the article is organized as follows: Section 2 provides a literature review and develops several hypotheses. Section 3 introduces the data and methodology. Section 4 presents the empirical results and findings. Conclusions and discussions are presented in the final section.

2 | RESEARCH BACKGROUND AND HYPOTHESES

Current scientific research is entering the era of international collaboration (Adams, 2013), which plays a significant role in evolution of science (Coccia & Bozeman, 2016; Coccia & Wang, 2016). Across the globe, the percentage of internationally coauthored papers grew from 13.2% to 19.2% between 2000 and 2013; for each country, this increase in international coauthorship occurred during the same period (National Science Board, 2016). The rise of the global scientific network is changing the Atlantic axis (USA and Europe) pattern formed after World War II (Adams, 2012) and reshapes global scientific landscape (Royal Society, 2011). The globalization of science has been driven by factors such as international division of labor among researchers (Durkheim, 1997; Niosi & Bellon, 1994), open innovation paradigm (Chesbrough, 2003), global challenges (Royal Society, 2011), and progress in Transportation and Information Communication Technology (Adams & Loach, 2015).

Location analysis is a customary and traditional task in human and economic geography. Similarly, position analysis is the center of gravity of social network research. So, location is to human and economic geography what position is to network science. The embeddedness literature and social network theory suggest that actors are embedded in social relations and interactions that influence economic outputs (Borgatti, Mehra, Brass, & Labianca, 2009; Boschma, 2005; Granovetter, 1985; Polanyi, 1944; Wanzenbock & Piribauer, 2018). One of the most fundamental axioms in social network analysis is that “a node’s position in the network structure determines in part the opportunities and constraints that it encounters, and in this way plays an important role in a node’s outcomes” (Borgatti et al., 2009). By occupying a central position in the network, a node can easily access intangibly external resources (Tsai, 2000). The globalization of science increases the interconnectedness and interdependence of countries, which are actively or passively integrated into the global innovation network, and network ties can be seen as a channel through which nations assimilate external knowledge and exchange information. Therefore, national knowledge output is not only attributable to in-house R&D inputs but also depends in part on one’s position in the network structure. The goal of this study is to investigate the effect of network position (degree centrality, structural holes, and small-world quotient) on national knowledge production.
2.1 | Degree centrality and knowledge output

Centrality measures the importance of nodes in the network, which is one of the most important structural features in network analysis (Freeman, 1979). Network centrality includes degree, betweenness, closeness, and eigenvector centrality, which are measured from different perspectives (Dong & Yang, 2016; Freeman, 1979). Different types of centrality play different roles in knowledge diffusion. Salman and Saives (2005) indicate that centrality variables tend to be correlated to one another, which could result in a multicollinearity of variables. Therefore, we choose degree centrality, which is widely discussed (Coffano, Foray, & Pezzoni, 2017; Mitze & Strotebeck, 2018; Tsai, 2001; Wang, Rodan, Fruin, & Xu, 2014). Degree centrality is a classic measure of network position, which refers to the number of other nodes directly connected to one node (Freeman, 1979). The greater a node’s degree centrality, the more innovation outputs are likely to be produced. First, a node with high degree centrality often occupies a central position and has an information advantage (Gilsing, Nooteboom, Vanhaverbeke, Duysters, & Oord, 2008). Specifically, these nodes are likely to have access to, share, and use desired, complementary, and heterogeneous resources that are critical to innovation activities (Borgatti, 2005; Tsai, 2001). Second, the reliability of networks directly depends on their vulnerability to disruptive incidents, ranging in severity from random breakdowns to intentional attacks (Bell, Kanturska, Schmocker, & Fonzone, 2008; Cohen, Erez, Ben‐Avraham, & Havlin, 2001, 2000). A node with high degree centrality has more direct partners than others, which means that it has more alternative channels and means of linking with indirect nodes to increase the stability of external resources. Third, as innovation generally arises from combining or recombining existing knowledge elements (Fleming, 2001; Schumpeter, 1934), knowledge diversity will enhance the possibility of combination or recombination (Coffano et al., 2017; Wanzenbock & Piribauer, 2018). A country with high degree centrality has multiple information types, which can be recombined through a unique and novel mode. Consequently, we advance the following hypothesis:

\[ H1: \text{A national degree centrality in the scientific collaboration network has a positive effect on its future knowledge output.} \]

2.2 | Structural holes and knowledge output

Different from degree centrality, which emphasizes direct ties, structural holes highlight nonrepeated ties. According to Burt’s structural hole theory, structural holes refer to the nonredundant links between two actors, that is, a specific node’s two neighbor nodes are disconnected from each other (Burt, 1992), and are a strategic and influential position in the network (Ahuja, 2000; Belso-Martinez, Diez-Vial, Lopez-Sanchez, & Mateu-Garcia, 2018). In a collaboration network, it is impossible for all node pairs to have direct ties, except for a globally coupled network. Therefore, structural holes are a universal phenomenon. Scott and Carrington (2011) argue that structural holes are able to play the role of a bridge or intermediary. Hence, the concept of structural holes is similar with brokerage or gatekeeper (Araújo et al., 2018; Belso-Martinez et al., 2018; Breschi & Lenzi, 2015). As an actor’s structural holes increase, its knowledge creation will improve for three reasons. First, actors spanning structural holes have advantages in accessing fresh information and controlling information diffusion (Burt, 1992; Granovetter, 1973; Zaheer & Soda, 2009) and can access novel information from remote partners and mediate the flow of knowledge and information between actors (Burt, 1992). Second, an actor maintaining a number of ties to many other actors is expensive to maintain (Goyal, 2007; Zaheer & Bell, 2005); when the number of ties reaches a threshold, information flows between actors
will be redundant and valueless (Leiponen & Helfat, 2010). Third, an actor bridging structural holes can receive fewer constraints and enjoy more autonomy in decision-making (Burt, 1992; Shipilov & Li, 2008; Wang et al., 2014). Thus, the actor can be free from the restrictions of its partners (Cheon, Choi, Kim, & Kwak, 2015), such as opinion leaders and prevailing cognitive schemes that may hinder the willingness and effectiveness of actors to continue innovation (Janis, 1972; Wang et al., 2014). In short, an actor with rich structural holes is more likely to acquire novel and heterogeneous information, have nonredundant ties and enjoy autonomy benefits, which will enhance an actor’s innovation performance.

**H2**: A national structural hole in the scientific collaboration network has a positive effect on its knowledge output.

### 2.3 Small-world and knowledge output

According to Watts and Strogatz (1998), small-world networks simultaneously possess a high clustering coefficient and a short characteristic path length. The clustering coefficient refers to the possibility of a specific node’s neighbor nodes connecting to each other, and characteristic path length is defined as the average number of edges along the shortest paths between all pairs of actors (Uzzi & Spiro, 2005; Watts & Strogatz, 1998). The small-world quotient is used to measure the degree of the network’s small-world nature and calculated by the clustering coefficient divided by the path length (Chen & Guan, 2010; He & Fallah, 2014; Uzzi & Spiro, 2005; Zhang, Guan, & Liu, 2014). Thus, it tends to be seen as the interaction term for the clustering coefficient and path length (Fleming et al., 2007; Schilling & Phelps, 2007). Numerous empirical studies suggest that the small-world property will enhance actors’ knowledge creation for two reasons. First, the collaboration network as a social network is built on friendship, trust, and past interaction. Moreover, actors are inclined to connect to partners of partners; such triangle relationships are called closure or triadic closure (Ter Wal, 2014). The high clustering and dense ties are likely to share resources, curb opportunism, disseminate risk, promote trust, and facilitate information transmission in the clustered network (Granovetter, 1985; Hung & Wang, 2010; Uzzi & Spiro, 2005). Second, information from remote partners may be distorted by intermediaries during the process of information diffusion (Araújo et al., 2018; Schilling & Phelps, 2007), and the validity of information decays with increasing distance or path length. Moreover, short path length can enhance information transmission efficiency, reduce the cost of collaboration, and expose actors to new information (Fleming et al., 2007; Gulati, Sytch, & Tatarynowicz, 2012). In short, small-world structure increases trust, reduces the distance between actors, and efficiently and effectively facilitates information transfer.

**H3**: A national small-world quotient in the scientific collaboration network has a positive effect on its knowledge output.

### 3 DATA AND METHOD

#### 3.1 Data

Most internationally coauthored papers are found in Elsevier’s Scopus database and Clarivate Analytics’s WoS. However, unless otherwise noted, this study analyzes data on international
collaboration from the WoS. Because the WoS includes approximately 20,000 high-quality scholarly journals published worldwide and is an ideal resource for the study of international collaboration in science. Moreover, it is often used in empirical studies (Cassi, Morrison, & Wal, 2012; Hoekman et al., 2009; Li & Phelps, 2017), which means that our selection is appropriate. In order to provide a wide-ranging snapshot of global scientific collaboration, our data cover all disciplines, all citation indexes (SCI-E, SSCI and A&HCI) and all document types (such as article, meeting abstract, editorial material, and review). In this paper, the WoS is extracted to retrieve data according to the following steps: (1) Extract each country’s annually collaborative partnerships between 2000 and 2015 from the Web of Science Core Citation database and save them as a TXT text file. (2) Construct a country-by-country association matrix using C++ program, where diagonal cell values denote the number of publications from country i in year t and other cells denote the number of copublications between country i and country j in year t. Since knowledge production is a cumulative and path-dependent process (Heimeriks & Boschma, 2014), we take a 5-year moving window to construct dynamic collaboration networks (Breschi & Lenzi, 2016; Cassi et al., 2012; Fleming et al., 2007). More specifically, a 5-year moving window procedure is applied to reconstruct collaboration data, resulting in a total of 11 symmetrically undirected and weighted networks. As explained below, R&D data are regarded as the control variables; otherwise, the estimation results may be biased. Given the availability of R&D data, our final networks contain 60 countries. In addition, as the number of copublications between countries shows significant differences, we prune those networks. Specifically, edges with more than 100 copublications are retained, while others are dropped. Alternate threshold values have little effect on the panel estimates, such as 20 and 50.

3.2 | Variables and statistical models

Although science does not necessarily lead to the publication of research papers, publication-based measures are suitable and widely accepted as proxies of knowledge output (Cantner & Rake, 2014; Guan, Zuo, Chen, & Yam, 2016; Royal Society, 2011). We use the number of publications as the measure of national knowledge output (NKO). A 5-year lags is introduced to minimize the possible problems of endogeneity and reverse causality in our model. More specifically, the control variables are measured by year t, the explanatory variables are calculated between year t and t + 4 (a 5-year moving window) and the dependent variable is measured by year t + 5. Taking the year of 2000, for example, the R&D inputs are from 2000, the network indicators correspond to data from 2000 to 2004, and the knowledge output is the number of papers published in 2005.

3.2.1 | Degree centrality (DC)

Nodal degree centrality indicates the number of edges directly connected with node i (Freeman, 1979). In an international scientific collaboration network, degree centrality is the number of countries that country i collaborates with

\[ DC_i = \sum_{j=1}^{N} a_{ij} \]  

(1)

where \( a_{ij} \) represents an adjacency matrix, \( a_{ij} = 1 \) when a link exists between country i and country j, and \( a_{ij} = 0 \) otherwise.
### 3.2.2 | Structural holes (SH)

Structural holes are measured based on network constraint (Burt, 1992). First, we calculate a dyadic constraint $C_{ij}$ between $i$ and $j$ with the formula below:

$$C_{ij} = \left( p_{ij} + \sum_{q \neq i, q \neq j} p_{iq}p_{qj} \right)^2$$  \hspace{1cm} (2)

where $q$ is the number of third-party countries to which both $i$ and $j$ are connected and $p_{ij}$ is the number of the focal country $i$’s network contacting with $j$ ($p_{iq}$ and $p_{qj}$ are defined analogously). For example, if $i$ is tied to three countries with equal strength, then $p_{ij}$ is $1/3$.

The next step is to aggregate constraint measure $C_i$ for country $i$:

$$C_i = \sum_j C_{ij}$$  \hspace{1cm} (3)

To estimate SH$_i$, we subtract $C_i$ from 2 (Lee, 2010) to represent the extent to which countries tied to a focal country $i$ are disconnected:

$$SH_i = 2 - C_i$$  \hspace{1cm} (4)

### 3.2.3 | Small-world quotient (SW)

Small-world networks simultaneously consider the clustering coefficient and path length (Watts & Strogatz, 1998). First, the clustering coefficient ($C$) measures the probability of a specific node’s neighbor nodes, which are connected to each other:

$$C_i = \frac{E_i}{K_i(K_i-1)/2}$$  \hspace{1cm} (5)

where $E_i$ is the actual number of edges among node $i$’s neighbor nodes and $K_i$ is the number of node $i$’s neighbor nodes.

Second, characteristic path length (PL) measures the average number of edges along the shortest path for node $i$ to reach other nodes, written as

$$PL_i = \frac{\sum_j d_{ij}}{n}$$  \hspace{1cm} (6)

where $d_{ij}$ is the number of edges for the shortest path between $i$ and $j$ and $n$ is the number of nodes $i$ directly and indirectly reaching them.

To estimate SW$_i$, we calculate the small-world measure as the clustering coefficient divided by characteristic path length (Shi & Guan, 2016):

$$SW_i = C_i/PL_i$$  \hspace{1cm} (7)

where $C_i$ is node $i$’s clustering coefficient and $PL_i$ is node $i$’s characteristic path length.
Taking into account the impact of in-house R&D inputs, we control several variables in the empirical models, such as RDE, research and development personnel (RDR) and economic development level (EDL). First, we use gross expenditure on R&D (% of GDP) as a proxy for spending on science. Second, the number of full-time equivalent researchers is the proxy used for human capital on science, which is measured by the logarithm of R&D researchers (Ln R&D researchers) in each country (researchers per million people). Third, we employ the World Bank income group to measure the EDL of a country. According to the World Bank, the analytical classifications for country incomes are (1) low income, (2) lower middle income, (3) upper middle income, and (4) high income.

As discussed above, Table 1 lists and describes the dependent, explanatory, and control variables in the econometric models.

The dependent variable, the number of publications, is count data. Because the conditional variance of the dependent variable obviously exceeds its conditional mean, namely existing overdispersion, we employ a negative binomial model instead of the Poisson model in our analysis (Araújo et al., 2018; Breschi & Lenzi, 2015; He & Fallah, 2014; Miguelez & Moreno, 2013; Ponds et al., 2010). In addition, the Hausman test rejects random effects specification. Therefore, we use fixed effect negative binomial models to explore our hypotheses.

4 | RESULTS

4.1 | Descriptive analysis

As shown in Figure 1, the annual number of scientific publications in major countries shows a steady increase in the period from 2000 to 2015. Among them, China shows remarkable growth and is becoming a science superpower. In 2000, China publishes 31,959 papers and ranks eighth. In 2011, China publishes 173,643 papers, second only to the USA. In 2015, China publishes 306,831 papers, 50% more than the UK. Compared with the number of China’s publications in 2000, it increases

| Table 1 Variable descriptions and data sources |
|-----------------|-----------------|-----------------|
| Variable name   | Description     | Source          |
| Dependent variable | NKO             | The number of publications in t + 5 year | Web of science |
| Explanatory variables | Degree centrality | The number of partners to which a focal country is directly connected | / |
|                  | Structural holes | The nonredundancy among ties in a focal country’s ego network is measured by network constraint | / |
|                  | Small-world quotient | The ratio of clustering coefficient and characteristic path length | / |
| Control variables | RDE             | The ratio of gross expenditure on R&D and GDP (% of GDP) | World bank |
|                  | RDR             | Log of R&D researchers in a focal country | World bank |
|                  | EDL             | Dummy variable, the economic development level of a focal country | World bank |
by 274,872 in 2015 and is consistent with the significant exponential law of $y = 28,403e^{0.1523x}$ ($R^2 = 0.992$). China’s scientific growth averages 16.28% per year between 2000 and 2015. In the same period, the world’s annual scientific publications growth rate is 4.6%, and the United States’ is 3.1%. It is worth mentioning that the USA has ranked first in terms of publications, maintaining a large gap with other countries, and its status as a science superpower is ongoing. In 2015, the United States’ scientific output reaches 587,353, producing 27.02% of global publications, about twice as much as that of China. European countries are relatively stable: the annual growth rate of the UK is 3.53%, Germany’s is 3.60%, France’s is 3.39%, and Italy’s is 5.79%. In Asia Pacific, Australia had an annual growth rate of 7.8%, and its growth trend is significant as well. However, Japan represents low-speed growth, with only 2.23% per year.

Figure 2 shows the evolution of the spatial structure of the international scientific collaboration network. The nodes represent countries, whose sizes are measured by weighted degree centrality, which is the sum of the article counts for each country’s tie. The edges denote the collaborative relationships between two countries, with the thickness of the edges scaled to show absolute strength, i.e., the number of copublications. As noted earlier, edges with no fewer than 100 copublications are retained; others are removed. In 2000–2004, the international scientific collaboration network is dominated by G7 countries, which includes the European Union (UK, Germany, Italy, and France), North America (USA, Canada), and Japan. The six largest collaborations occur between the USA and the other six countries. In addition, the total number of lines is 793, and the value of the network density is 0.45, which indicates that the whole network is weakly connected.

It is observed obviously that the global scientific landscape in 2010–2014 has undergone fundamental changes. The international cooperation once led by Europe and the United States is gradually replaced by the three-polar world including Europe, North America, and Asia Pacific region. Bilateral partnerships between the USA and China outnumber all other international pairings. The number of knowledge production in other emerging nations, such as South Korea, Brazil, Singapore, South Africa, and Turkey, is rapidly rising. Traditional scientific superpowers and emerging scientific nations coexist and are reshaping the scientific landscape. Moreover, it is also seen that international
collaboration is not severely constrained by geographical distance and the main links occur between hub countries. In addition, the preferential attachment phenomenon is observed. Countries with low output are more likely to establish ties with countries with high output instead of collaborating with other countries that are geographically close (Guan, Zhang, & Yan, 2015). Finally, the total number of lines is 1,340, and the value of the network density is 0.76, which means that international collaboration is more frequent and closer than before.

4.2 Regression analysis

Table 2 presents descriptive statistics and a correlation matrix. In our study, countries publish a mean of 32,859.52 papers per year, with a standard deviation of 72,832.04 papers, and the latter is more than twice the former, which reflects the heterogeneity of national innovation performance. By contrast, the standard deviations of explanatory and control variables are lower than their means. The average variance inflation factor is 2.71, and every variance inflation factor value is lower than 5, which suggests that multicollinearity is not significant in this paper.

FIGURE 2   Spatial patterns of international scientific collaboration network: (a) 2000–2004, (b) 2010–2014
Table 3 displays the results of fixed effect negative binomial models on national innovation performance. Model 1 in Table 3 is the basic model, which shows the estimated results of the control variables. Models 2, 3, and 4 are constructed based on Model 1, while to each are added degree centrality, structure holes, and small-world quotient, respectively. Model 5 presents the full model, which includes all control variables and explanatory variables. Not surprisingly, R&D expenditure, the number of researchers and the EDL present positive and statistically significant effects.

Table 3 displays the results of fixed effect negative binomial models on national innovation performance. Model 1 in Table 3 is the basic model, which shows the estimated results of the control variables. Models 2, 3, and 4 are constructed based on Model 1, while to each are added degree centrality, structure holes, and small-world quotient, respectively. Model 5 presents the full model, which includes all control variables and explanatory variables. Not surprisingly, R&D expenditure, the number of researchers and the EDL present positive and statistically significant effects.

Table 2

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKO</td>
<td>32859.520</td>
<td>72832.040</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree centrality</td>
<td>35.170</td>
<td>16.639</td>
<td>0.437***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural holes</td>
<td>1.718</td>
<td>0.115</td>
<td>0.196***</td>
<td>0.665***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small-world quotient</td>
<td>0.640</td>
<td>0.106</td>
<td>0.042</td>
<td>0.653***</td>
<td>0.712***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R&amp;D expenditure</td>
<td>1.339</td>
<td>1.014</td>
<td>0.336***</td>
<td>0.533***</td>
<td>0.258***</td>
<td>0.242***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln (R&amp;D researchers)</td>
<td>7.339</td>
<td>1.169</td>
<td>0.216***</td>
<td>0.519***</td>
<td>0.427***</td>
<td>0.324***</td>
<td>0.861***</td>
<td></td>
</tr>
<tr>
<td>Economic development</td>
<td>3.414</td>
<td>0.749</td>
<td>0.177***</td>
<td>0.425***</td>
<td>0.338***</td>
<td>0.231***</td>
<td>0.553***</td>
<td>0.677***</td>
</tr>
</tbody>
</table>

***p < 0.01, **p < 0.05, *p < 0.1.

Table 3

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;D expenditure</td>
<td>0.1412***</td>
<td>0.0736***</td>
<td>0.0703**</td>
<td>0.0228</td>
<td>0.0537**</td>
</tr>
<tr>
<td></td>
<td>(0.0375)</td>
<td>(0.0281)</td>
<td>(0.0288)</td>
<td>(0.0311)</td>
<td>(0.0272)</td>
</tr>
<tr>
<td>Ln (R&amp;D researchers)</td>
<td>0.5092***</td>
<td>0.3061***</td>
<td>0.3253***</td>
<td>0.4034***</td>
<td>0.2821***</td>
</tr>
<tr>
<td></td>
<td>(0.0515)</td>
<td>(0.0388)</td>
<td>(0.0372)</td>
<td>(0.0394)</td>
<td>(0.0359)</td>
</tr>
<tr>
<td>Economic development</td>
<td>0.2604***</td>
<td>0.1122***</td>
<td>0.0843***</td>
<td>0.1827***</td>
<td>0.1002***</td>
</tr>
<tr>
<td></td>
<td>(0.0318)</td>
<td>(0.0220)</td>
<td>(0.0223)</td>
<td>(0.0240)</td>
<td>(0.0210)</td>
</tr>
<tr>
<td>Degree centrality</td>
<td>0.0223***</td>
<td></td>
<td></td>
<td></td>
<td>0.0133***</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td></td>
<td></td>
<td></td>
<td>(0.0014)</td>
</tr>
<tr>
<td>Structure holes</td>
<td></td>
<td>3.4611***</td>
<td></td>
<td></td>
<td>0.9000***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.1402)</td>
<td></td>
<td></td>
<td>(0.2278)</td>
</tr>
<tr>
<td>Small-world quotient</td>
<td></td>
<td></td>
<td>1.6155***</td>
<td></td>
<td>0.4628***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0816)</td>
<td></td>
<td>(0.0928)</td>
</tr>
<tr>
<td>Constant</td>
<td>−1.5945***</td>
<td>0.5336**</td>
<td>−4.7531***</td>
<td>−0.8376***</td>
<td>−0.6358*</td>
</tr>
<tr>
<td></td>
<td>(0.3302)</td>
<td>(0.2539)</td>
<td>(0.2856)</td>
<td>(0.2493)</td>
<td>(0.3819)</td>
</tr>
<tr>
<td>Observations</td>
<td>660</td>
<td>660</td>
<td>660</td>
<td>660</td>
<td>660</td>
</tr>
<tr>
<td>Number of country</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>−5319.3086</td>
<td>−5067.9675</td>
<td>−5095.8803</td>
<td>−5140.4813</td>
<td>−5035.5106</td>
</tr>
<tr>
<td>Wald $X^2$</td>
<td>521.90</td>
<td>1859.98</td>
<td>1600.01</td>
<td>1629.26</td>
<td>2264.34</td>
</tr>
</tbody>
</table>

Note. Standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.
Hypothesis 1 predicts that a country’s degree centrality in an international scientific collaboration network is positively related to its future knowledge output. As shown in Models 2 and 5 in Table 3, the coefficient of the degree centrality of a focal country is positive and statistically significant at \( p < 0.01 \). Countries with high degree centrality tend to publish more papers than those with low degree centrality. Therefore, a country can benefit from its number of direct partners in the collaboration network. This not only corresponds with the findings of Guan et al. (2016), who use country-level coauthored papers data, but also with those of Coffano et al. (2017), who use inventors’ collaboration networks in the Swiss medical devices sector. The increased number of direct ties of a country in an international scientific collaboration network is good for obtaining desired strategic resources and improving the stability and diversity of external knowledge sources. The increase of one unit in degree centrality corresponds to a 24.76% \((= e^{0.0133*16.639} - 1)*100\) increase in knowledge output, with all other variables held constant. Hence, Hypothesis 1 is confirmed.

Hypothesis 2 argues that a country with more structural holes in an international scientific collaboration network tends to produce better innovation performance. The coefficient of the structural holes is positive and statistically significant at \( p < 0.01 \) in Models 3 and 5. A country collaborating with disconnected countries tends to increase knowledge output. This result is in line with comparable prior studies that employ interfirm collaboration network in the US microprocessor manufacturer (Wang et al., 2014) and metropolitan coinvention networks in US (Breschi & Lenzi, 2015). An actor with rich structural holes is more likely to acquire novel and heterogeneous information, have nonredundant ties, and enjoy autonomy benefits, which will enhance innovation performance. A one-unit increase in structural holes corresponds to a 10.9% \((= e^{0.9*0.115} - 1)*100\) increase in knowledge performance in the full model. Therefore, Hypothesis 2 is supported by empirical results.

Hypothesis 3 states that a country’s small-world quotient in an international scientific collaboration network will facilitate its innovation performance. The reported regression coefficients of Models 4 and 5 in Table 3 are positive and statistically significant at \( p < 0.01 \). A country with a high clustering coefficient and short characteristic path length in the collaboration network can have more knowledge output than others. Our empirical result is consistent with various empirical studies for the patent collaboration networks of 16 countries (Chen & Guan, 2010), interfirm collaboration networks on the basis of 11 high-technology manufacturing industries alliance (Schilling & Phelps, 2007), and hi-tech industry metropolitan clusters using inventor collaboration networks (He & Fallah, 2014). The small-world structure increases trust and reduces distance between actors, and facilitates efficient and effective information transfer. A one standard deviation increase in the variable tends to raise the rate of paper publication by a factor of 0.05 \((= e^{0.4628*0.106} - 1)*100\) in Model 5. Thus, our findings support Hypothesis 3.

In unreported results, we also try to consider the quadratic term of the explanatory variables, but these are not statistically significant and robust.

### 4.3 Robustness analysis

To further test our hypothesis, a robust test was conducted. The total number of national papers can be divided into two parts: internationally coauthored papers and domestic papers. In a sense, internationally collaborative outputs (ICO) are more easily influenced by network structure. In addition, international copublications are cited relatively more often than purely domestic publications (Adams, 2013). As a consequence, an additional negative binomial regression with ICO is designed to test the robustness of our results. The dependent variable is measured by the number of internationally coauthored papers. The control and explanatory variables are the same as regression analysis. As demonstrated in Table 4, degree centrality, structural holes, and small-world quotient exhibit positive and statistically
significant impacts on ICO. Therefore, our previous findings from total national publications are supported by international copublications. Moreover, it is also evidenced that network structure has a greater impact on the latter.

5 | DISCUSSION AND CONCLUSIONS

Despite social networks are important mechanism of knowledge spillovers, there is a scarcity of analysis on the influence of network structure on knowledge production. Using data on international copublications from the Web of Science core database, this paper constructs eleven country-level collaboration networks between 2000 and 2015 and investigates the effect of degree centrality, structural holes, and small-world quotient on national knowledge output. This article provides empirical evidence to support our research hypothesis. The regression results show that the coefficients of the three network properties are positive and statistically significant in panel data estimates, which means that occupying a central and strategic position in the network is instrumental for national knowledge production.

Our empirical analysis sheds some novel and interesting light on knowledge production (Griliches, 1979) and knowledge spillovers (Jaffe, Trajtenberg, & Henderson, 1993). First, this study enriches the burgeoning body of literature on the determinants of knowledge output or innovation performance. The traditional knowledge production research explain knowledge outcomes as a function of tangible inputs (RDE or human capital), while a growing number of literature look to intangible or soft factors for explanations, such as culture and institutions, 3Ts of talent, technology, and tolerance (Florida,

| TABLE 4 | Results of conditional fixed effect negative binomial models of ICO |
|----------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Variables | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
| R&D expenditure | 0.1562*** | 0.0663* | 0.0846*** | −0.0105 | 0.0303 |
| Ln (R&D researchers) | 0.4391*** | 0.3010*** | 0.3022*** | 0.4140*** | 0.2979*** |
| Economic development | 0.2453*** | 0.0550** | 0.0196 | 0.1389*** | 0.0551*** |
| Degree centrality | 0.0304*** | 0.0131*** | (0.0010) | (0.0015) |
| Structure holes | 4.9551*** | 2.5085*** | (0.1543) | 0.9874*** | (0.2459) |
| Small-world quotient | −1.5962*** | 0.1225 | −7.2059*** | −1.3244*** | −1.2667*** |
| Constant | (0.3864) | (0.3018) | (0.3283) | (0.2527) | (0.4001) |
| Observations | 660 | 660 | 660 | 660 | 660 |
| Number of country | 60 | 60 | 60 | 60 | 60 |
| Log likelihood | −4993.8321 | −4687.7989 | −4685.4621 | −4661.0485 | −4555.3731 |
| Wald χ² | 304.16 | 1712.36 | 1790.34 | 2541.64 | 3527.00 |

Note. Standard errors are shown in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.
2002), creativity (Marrocu & Paci, 2012), psychology (Lee, 2017). However, the most important intangible asset may be social capital (Dettori, Marrocu, & Paci, 2012), which is the resource obtained from the interactions between actors in social networks (Burt, 1992; Huber, 2009; Lin, 2002). As Burt (2005) put it, the rate of return on an actor’s tangible inputs depends on their social capital (network position). Second, the article increases the growing literature on relational spillovers effects. It is noted that knowledge spillovers not only occur in the local scale, but also take place at a global scale (Bathelt et al., 2004); not only occur through informal knowledge exchange, but also take place through formal networks of research collaboration (Ponds et al., 2010); not only represent spatial spillovers, but also perform relational spillovers (Maggioini et al., 2007). In the context of the globalization, informatization, and networking, the determinants of knowledge spillovers has undergone fundamental changes, most notably the ups of relational and capability proximity and the downs of geographical and cultural proximity (Gui et al., 2018; Montobbio, Primi, & Sterzi, 2015; Ter Wal, 2014). Formal collaboration networks will be more important than informal knowledge exchange, because the former can occur over longer geographical distances (Ponds et al., 2010) and fresh and novel information from remote partners declines the risk of technological lock-in (Crespo et al., 2014; He & Fallah, 2014).

This paper provides important implications for policymakers. First, it is found that the structure of the international scientific collaboration network has a positive and significant effect on national research output, and countries can benefit from a central and influential position in the network. Therefore, internationally collaborative science should be encouraged, supported, and facilitated (Royal Society, 2011). Given the globalization of science and technology, it is particularly important for a country to access external resources. These network relationships can act as channels that create, diffuse, absorb, exploit and share ideas, resources, knowledge, and information (Phelps, Heidl, & Wadhwa, 2012). Second, intercountry scientific collaborations seem to be a top-down collaboration. In essence, international research collaborations are an aggregation of interindividual- and interinstitutional-level collaboration, which are mainly driven by the bottom-up strength. National policies should incentivize and stimulate researchers and institutions (e.g., universities and firms) to participate in international collaboration networks and help universities and researchers create and sustain flourishing partnerships (Adams, 2013). In terms of the individual level, Lee and Bozeman (2005) find that research grants have significant effect on scientific productivity. The goal of Europe 2020 strategy is to build an integrated European Research Area and encourage international collaboration. For example, the EU’s “research budget” a significant portion of the funding is directly earmarked for international projects.

This article is meaningful, but has a few limitations. First, we concentrate only on intercountry collaboration networks. However, knowledge spillovers are doubly embedded in formal research collaboration and informal social networks. The latter may act as an invisible college (Price, 1986). Second, we investigate the impacts of network position (degree centrality, structural holes, and small-world quotient) on national knowledge output, but no attention is paid to the question that what can influence a country’s position in the network structure. Third, knowledge spillovers take place at different geographical scales. The empirical framework of this paper can be adopted at the meso-level scales, such as the city or region level.

ACKNOWLEDGMENTS

The authors would like to acknowledge funding from the National Natural Science Foundation of China (No. 41571123; No.41471108), Key Project of Soft Science of Shanghai (No.17692103600), and Shanghai Pujiang Program (No.17PJJC030).
REFERENCES


---

**How to cite this article:** Gui Q, Liu C, Du D. Does network position foster knowledge production? Evidence from international scientific collaboration network. *Growth and Change*. 2018;00:1–18. [https://doi.org/10.1111/grow.12263](https://doi.org/10.1111/grow.12263)