Networks and Spatial Economics (2022) 22:65–96
https://doi.org/10.1007/s11067-021-09556-3

The Network Structure of Innovation Networks

Shixun Wang[1] · **Lihong Yang**[2,3,4]

Accepted: 8 September 2021 / Published online: 26 January 2022

Abstract
We develop a model of R&D networks in which firms can choose to invest in R&D or to establish network links with other firms to absorb R&D spillovers in innovation networks. We characterize the network structure of firms' strategic R&D decisions in different types of spatial equilibrium. We further extend the theoretical framework to address empirical implications for the industry structure under cooperative R&D, firms' location choice, spatial agglomeration, and social capital allocation.

Keywords Innovation networks · Spillover · Strategic R&D · Diffusion · Equilibrium

1 Introduction

Technological progress builds upon itself, expanding innovation from one firm propelling future work for other linked firms (Acemoglu et al. 2016). Intensified global competition means that, at present, no firm can remain competitive by relying entirely on its internal resources and capabilities. Innovation alliances and partnerships have become widespread, especially in high-tech industries characterized by rapid knowledge spillovers, such as the computer, chemical, and pharmaceutical industries (Hagedoorn 2002; Roijakkers and Hagedoorn 2006). Through such collaborations in innovation networks, firms may generate technological spillovers directly for their cooperative partners and propel economic growth and long-term well-being.

✉ Lihong Yang
lihong.yang@ruc.edu.cn

Shixun Wang
wangshixun@ruc.edu.cn

1 School of Economics and Management, Fuzhou University, Fuzhou, China

2 School of Economics, Nanjing Audit University, Nanjing, China

3 International Business School, Zhejiang University, Hangzhou, China

4 National Academy of Development Strategy, Renmin University of China, Beijing, China

A growing body of literature studies different network structures and new implications of network economy based on theoretical foundations in the field of industrial organization by recognizing the fact that, in a network economy, firms adopt a "local strategy" rather than "global strategy." When the player is in a different position in the network, its response to other players is different. At the same time, it also has different effects on others (Vives 2009; Aoyagi 2018; Manea 2018). Compared with most existing research that assumes that the network structure is exogenously given and independent of the firms' behavior,[1] a firm's network links and their market behavior usually interact with one another in a network economy.[2] Concerning R&D, to reconcile the interaction between network structure and firms' behavior, the actions of firms in this model include market behavior (investment in R&D) and network behavior (connections established with other firms), where the network structure is not an exogenous parameter but part of the firm's strategy in equilibrium. In the context of technological innovation, the firm network reflects the technological spillovers. A firm can invest in R&D or connect with other firms and absorb others' R&D investment (R&D spillovers). Then, the firm's actions include R&D investment and with which firms to connect.

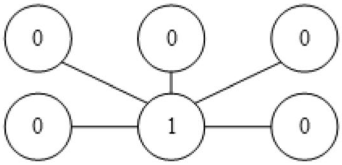
We develop a theoretical network game to analyze the network structure by incorporating endogenous R&D spillovers and diffusion effects in innovation networks. We find that network structures in equilibrium can be summarized into two main types. As shown in the following figure, the first type is a "core" network. Some firms are at the center of the network and play a significant R&D role as an innovator. Other firms on the periphery establish links with these firms as imitators. In this network, the degree of innovation is relatively high, taking a significant proportion of the network's aggregate innovation degree and indicating a high-centrality network. The second type is a "multicluster" network. Some firms play a considerable R&D role as innovators, while other firms establish contacts as imitators, forming multiple clusters. In this network, the degree of innovation is still high. Still, it occupies a minor aggregate innovation degree than the core network, indicating that the multicluster network is a lower-centrality network. Of the two network structures, the core network appears when R&D resources are complementary, while the multicluster network appears in the context of substitute R&D resources.

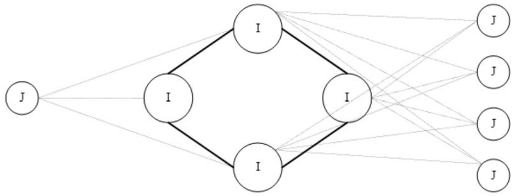
Although the centrality of the two types of networks is different, their centrality is still high compared to other networks, such as a minimum connected graph. In other words, even if the diffusion of R&D spillovers is far-reaching, which means that firms can absorb not only the R&D investments of their neighbors but also the R&D investments of neighbors' neighbors, and so forth, the network structure in equilibrium will limit the spillover effect within a narrow range, i.e., the distance between

[1] For example, firms may play a location choice game in which the network is formed by geographical location and the geographical location of the firm is given.

[2] In the case of the location choice game, a firm may adjust its network structure by choosing its location and which market to enter, which in turn affect the market strategy in equilibrium. Firms that do not have a cost advantage can avoid entering a market with many competitors and choose a relatively remote place.

any two firms is 1 or 2. If we regard the network as a geographic network, the results imply that the world is not "flat" and that spatial agglomeration in R&D still exists, even in the Internet era.





In the theoretical model, we abstractly analyze network characteristics in equilibrium without specific economic interpretations. In the following section, we consider the implication of the findings. The first application relates to cooperative R&D, where technology spillover depends mainly on disseminating knowledge and information. We find that a reduction in technology spillover costs reduces the cost of imitation and facilitates cooperative R&D. Under these two effects, the number of innovators will eventually rise, and the number of imitators will decrease. Therefore, reducing technology spillover costs due to more convenient propagation conditions reduces intellectual property protection difficulty. The second application relates to spatial agglomeration, where R&D spillovers are affected by the geographical distance between firms. With the development of information technology, technology spillovers are no longer limited to a particular geographical distance but can spread to distant places. Will firms still form traditional spatial agglomerations? Based on the equilibrium results, we find that traditional spatial agglomeration persists, but a new cluster model, called "multicluster agglomeration," has arisen. The third application relates to centralization and decentralization concerning innovation and competition policy. The firm's location in the network will have different impacts on other firms and is a source of market power. In a more centralized network, a small number of firms tend to have greater market power. The opposite is true in networks with lower levels of centralization. We find that depending on the initial network structure and the degree of centralization, a change in research and development costs and technology spillover costs, among other facts, will have diametrically opposite effects on the degree of centralization.

This paper relates primarily to two areas of research. Previous studies have examined the relationship between network structure and information diffusion, innovation behavior, and innovation performance (Capello and Faggian 2005; Gilsing et al. 2008; Li and Tang 2019). Network structure belongs to exogenous parameters in these studies; that is, network structure affects information diffusion and dissemination and then affects innovation performance. Alternatively, according to Borgatti and Halgin (2011), it belongs to the research category of "network theory." This paper attempts to build another underlying mechanism: the network structure is an endogenous parameter, which will adjust according to the external environment and affect firms' innovation behavior and performance. This research, which regards network structure as an endogenous variable, belongs to the "theory of network," and previous studies pay less attention. However, some recent empirical studies suggest that when the innovation alliance faces the risk of information leakage to competitors, it will adjust the organizational form of innovation alliance (Ryu et al. 2018). This paper's research also echoes these findings, analyzing the formation mechanism and network structure characteristics more systematically.

Although social network-related research also focuses on network formation from a broader perspective, its analysis and conclusion should not be directly applied to innovation alliance, mainly because innovation activities are strategic and spillover. In the pioneering research of D'asprimont and Jacquemin (1988), when firms carry out innovation activities, they need to consider their direct benefits (such as cost reduction) and the impact of R&D spillover on competitors and partners. Therefore, when firms cope with different research joint ventures, product alliances, and supply chain partners, their strategic responses will be different (Kamien and Zhang 2000; Chu et al. 2018; Zhang et al. 2018; Federico et al. 2018). Therefore, in this study, forming an innovation alliance network that interacts with the R&D activities of firms also has strategic characteristics, which can be implied from the follow-up conclusion that this strategy is indeed the critical factor affecting network structure formation.

The second concerns the diffusion effect in the network. Some studies confirm that the scale-free structure or existence of highly central hubs in the network accelerates the transfer and diffusion of knowledge within the network (Qiao et al. 2019) and examine multiple factors of collaborative innovation networks that affect the level of knowledge transfer performance of firms (Xie et al. 2016). R&D spillovers are closely related to the diffusion effect in networks. More studies analyze the network structure when direct diffusion effects can only affect adjacent nodes (Calvó-Armengol 2009; Galeotti and Goyal 2010; Gagnon and Goyal 2017; Kireyev & Leonidov 2018).

This paper considers the network structure when the diffusion effect can influence all network nodes (after attenuation). We first take the supply chain as an example. There is a "bullwhip effect" in the supply chain, where low volatility in the final consumer's demand may cause considerable changes in the inventory and sales of upstream firms. When a retailer implements a supply chain management information system that improves consumer demand prediction accuracy, all firms in the supply chain will benefit.[3] Another issue concerns interdependencies among firms through financial tools (Elliott et al. 2014), where technology spillover may indirectly affect equity, credit, or other interest correlations. For example, Nissan and Renault form a stock-for-stock alliance in which the technology from Nissan may benefit the firms invested by Renault through the relationship.

The remainder of the paper is organized as follows. The second part describes the basic model. We analyze the network structure by establishing the endogenous network game model when the R&D spillover range is sufficiently large. The third part further analyses the equilibrium results and describes the conditions of different network structure features and their interaction with firms' behavior. The

[3] It is also required that the supply chain should be the "pull" type, indicating that production is determined by the order of downstream firms.

fourth part is the application. We accord the network with economic features and apply them to the firms' location choices and innovation alliances. The fifth part offers the conclusion and implications. In the appendix, we extend the model further to a standard two-stage R&D model and capture uncertainty and innovation spillovers.

2 Literature Review

2.1 Innovation and Network Structure

Many empirical studies have pointed out that network structure has a significant impact on innovation activities, but there are different views on the form and result of the impact. For example, Capello and Faggian (2005) document that firms' geographical agglomeration contributes to knowledge spillover and diffusion. Giuliani (2013) further confirms that with the dynamic evolution of firm agglomeration, knowledge exchange and technology reciprocity between firms contribute to forming new knowledge connections between firms in a dense network. In contrast, Rowley et al. (2000) claim that firms' substantial ties damage firms' performance in a highly interconnected alliance. Gilsing et al. (2008) found an inverted U-shaped relationship between innovation performance and innovation alliance centers.

Although there are different opinions on the relationship between network structure, especially network density and innovation activities, some recent empirical studies suggest that this complex influence may be related to the formation process of innovation alliance. For example, Ryu (2018) found that if a firm's partners are geographically adjacent to their competitors in the innovation alliance, the innovation alliance results may be leaked to the competitors to protect its knowledge employing equity investment and scope limitation. Aggarwal (2020) analyzes the indirect spillover effect of knowledge; the firms of innovation alliance can obtain resources from their partners and partners. However, due to capacity constraints, the use of such resources will be limited to a certain extent, resulting in resource congestion, which harms firms' innovation output.

It becomes necessary to theoretically analyze the formation mechanism of the innovation alliance network and its network structure to provide a micro foundation for the analysis of innovation alliance. The latest empirical research also implies that the network structure of innovation alliance is not "uniform." One firm group's network structure may differ from other network groups, which also has a heterogeneous impact on innovation performance. The model also helps to analyze further how different firm groups are formed and their network structure characteristics.

2.2 Innovation and The Theory of Network

According to Borgatti and Halgin (2011), network theory can be divided into two categories. The first is network theory: the mechanisms and processes that interact with network structures to yield cancer outcomes. In this kind of research, networks are regarded as exogenous parameters or independent variables, and researchers analyze the influence of networks from the individual or macro network perspectives (Jackson et al. 2017). Specific to innovation activities, this type of research is relatively more, including Capello and Faggian (2005), Gilsing et al. (2008), Li and Tang (2019).

The second research category is "theory of network," namely "the processes that determine why networks have the structures they do." In these studies, the network is regarded as an endogenous parameter or dependent variable, and a large part of the related literature focuses on the formation of social networks (Schilling 2001; Lee 2010; Jan et al. 2012). The main reason is that innovation activities have spillover effects and strategy, making firms' behavior characteristics different from the individuals in social networks and then making the formation of innovation networks unique.

It is worth mentioning that some studies focus on the regional agglomeration of innovation activities, which is also the manifestation of innovation network to a certain extent (Desmet and Rossi-Hansberg 2014; Tsiotas and Polyzos 2018). However, regional agglomeration only reflects the agglomeration characteristics of network structure but can not describe other network structure characteristics, such as multicluster, the central position of minority groups. Second, the analysis of the formation mechanism of regional agglomeration is generally only applicable to geographical networks but not to equity networks formed by equity investment or innovation alliance networks formed by R.J.V.

This study focuses on the "theory of network," a relatively less concerned area in previous research. Compared with the existing literature in this field, this paper highlights the spillover effect of individual behavior (innovation activities) and the influence of strategy in network formation and also describes the formation mechanism of various network structure characteristics.

2.3 Strategic Innovation

Most studies on R&D follow the theoretical framework in D'Aspremont & Jacquemin (1988). There are two stages in the R&D game. In stage 1, firms choose their R&D investment. In stage 2, firms compete with the Cournot game, and a higher R&D investment in stage 1 leads to a higher marginal cost saving. Kamien et al. (1992) extend the model to the case of research joint ventures and R&D cartels where firms can choose to cooperate or not cooperate in both stages.

Later studies using the R&D strategy follow two streams. The first considers endogenous technology spillover. In D'Aspremont & Jacquemin (1988) or Kamien et al. (1992), technology spillover is characterized by an exogenous parameter β, indicating that R&D investment can be absorbed by any firm with attenuation of β (i.e.,). In Kamien and Zhang (2000) and Amir et al. (2003), technology spillover is determined by the firm's strategy. For example, Kamien and Zhang (2000) extend the model to a three-stage game and introduce a parameter to capture the absorptive capacity. If the firm chooses an R&D approach with a lower δ_i (such as general technology), it can easily absorb technology spillover from other firms, while its technology is also easily absorbed by others. This study finds that technology spillover is affected by the choice of R&D approach, while the influence is still global. Once firm i chooses a certain δ_i, its absorptive capacity is identical to that of any other firm.

Some recent studies have identified differences in R&D decisions and welfare effects in supply chains, vertical constraints in retail firms, and horizontal mergers and acquisitions (Chu et al. 2018; Zhang et al. 2018; Federico et al. 2018). Although they are not directly related to technology spillover, these studies imply that the market structure and the relationships between firms have an essential effect on R&D behavior. We argue that the supply chain is one of the most important relationships between firms, but there are still various types of relationships in addition to it. For example, firms can cooperate in R&D and establish research joint ventures, as in Kamien et al. (1992), while several research joint ventures will compete.

To summarise, most studies on R&D games usually regard technology spillover as a part of a firm's strategy, but the spillover effect is global. In contrast, we follow the standard-setting in R&D games but introduce a network structure in which firms can choose whether to establish links with others and obtain technology spillover. Therefore, technology spillover is endogenous and has a local effect as a result. The network graph becomes crucial to capture the local effect, a unique feature for using a network model in this study. Recent studies also notice that firms' relationship has a significant influence but does not directly investigate how the relationship interacts with the firm's R&D strategy. This study considers the interaction between network structure and innovation strategy and discusses how a particular innovation alliance comes into being and how it interacts with the innovation strategy.[4]

Studies regarding networks confirm that the scale-free structure or existence of highly central hubs in the network accelerates the transfer and diffusion of knowledge within the network (Qiao et al. 2019) and examine multiple factors of collaborative innovation networks that affect the level of knowledge transfer performance of firms (Xie et al. 2016). Meanwhile, some studies also introduce network structure in the classical industrial organization model, such as Cournot competition (Bramoullé et al. 2014) and Bertrand competition (Aoyagi 2018). In these studies, the network and thus the position of firms in the network are exogenously determined.

Our study analyzes how an equilibrium network comes into being and its interaction with the R&D strategy. In other words, a network structure is part of a firm's strategy. For the methodology, we refer to Calvó-Armengol (2009), Galeotti and Goyal (2010), and Gagnon and Goyal (2017), who analyze the network structure with the diffusion effect. While the diffusion effect affects only adjacent nodes in these studies, we consider the case in which the diffusion effect will spread to the whole network.[5] Our study addresses the importance of interaction between R&D games and network economy and contributes to the classical R&D game by

[4] See a detailed discussion in 4.1 and 4.3.

[5] This setting is important for the results in Sect. 4.2 Spatial agglomeration of innovation network.

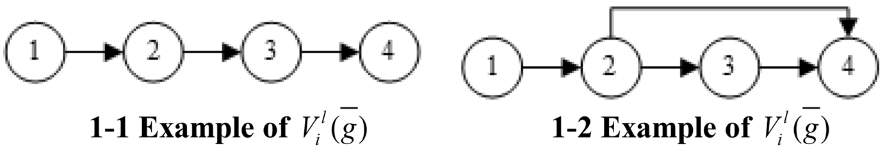


Fig. 1 Figs. 1-1 and 1-2 example of $V_i^l\ (\overline{g})$

introducing endogenous network structure. It also contributes to network economics by considering an endogenous network as well as technology spillover.

3 The Model

3.1 Setup

D'Aspremont & Jacquemin (1988) proposed a two-stage collaborative R&D model. In the first stage, the firm conducts R&D activities. In the second stage, the firm competes in the product market, and its marginal cost is affected by the R&D investment in the first stage. We include only the R&D behavior during the first stage in the basic model. We demonstrate in Appendix A that the basic model's conclusions can be extended to the standard two-stage model.

The set of firms is $V = \{1,\ 2, ..., n\}$. When making R&D decisions, a firm can directly invest x_i in research with unit cost c. The firm can also invest in technology spillovers to absorb R&D investments from other firms. We model R&D spillovers through the connections between firms, where $g_{ij} = 1$ indicates that firm *I* establishes a link with firm *j* and pays cost *k*. In most R&D models, R&D spillovers are identical for all firms (Kamien et al. 1992; Amir et al. 2003). However, some studies have shown that R&D spillovers are firm-specific (Gil Moultó et al. 2005; Cassiman and Veugelers 2002). Therefore, we use the relationship between firm *i* and firm *j* to model the technology spillover. In the adjacency matrix g, if there is $g_{ij} = 1$ or $g_{ji} = 1$, there is a direct R&D spillover between the two firms. If there is a path between firm *i* and firm *j*, there is an indirect R&D spillover between the two firms, which means that the R&D investment of firm *i* can be absorbed by firm j through the path.[6]

R&D spillovers can be either undirected or directed. An undirected R&D spillover indicates that if there is $g_{ij} = 1$, firm *i* and firm *j* have R&D spillovers with one another. For example, firm *i* completed the acquisition of firm j, so firm *i* took the initiative to establish a link with firm *j*. However, both firms can achieve R&D spillovers with the other. A directed R&D spillover indicates that if there is only $g_{ij} = 1$, then firm *i* can absorb the R&D investment of firm *j*, but firm j cannot absorb that of firm *i*. For example, if firm *i* reverse engineers the products of firm *j*, firm *i* can absorb an R&D spillover, while firm *j* cannot. In this paper, we focus on the case of

[6] If it exists $i,\ i_1, i_2, \ldots,\ i_k \in V$ such that $g_{i,i_1} = 1, \ldots, g_{ik,j} = 1$ and node *j*.

undirected R&D spillovers. Formally, regarding Galeotti and Goyal (2010) methods, the spillover effects of firm R&D are modeled as follows.

Let g be the adjacency matrix. If firm I establishes a link with firm j, we have $g_{ij}=1$, and $g_{ij}=0$ otherwise (specifically, $g_{ii}=0$). Let $\overline{g_{ij}} = max\{g_{ij}, g_{ji}\}$, and the corresponding graph is the undirected graph. Let $V_i^l(\overline{g})$ be the set of firms within distance l from firm I in this undirected graph. From the example in Fig.1–2, the distance between firm 1 and firm 4 is three, which means that firm 1 can be linked to firm 4 within three edges. Therefore, we have $V_1^3(\overline{g}) = \{4\}$. Note that there will be multiple paths between firm i and firm j and $V_i^l(\overline{g})$ indicates the shortest length of all paths. Figures 1–2 shows the distance between firm 1 and firm 4 is three. Therefore, we have. As the distance needed to make contact increases, so does the firm's R&D spillover. When the cardinality of set V is $|V| = n$, the distance between any two firms is less than n-1. Therefore, we define the R&D spillover as follows:

$$\sum_{l=1}^{n-1} \beta^{l-1} \sum_{j \in V_i^l(g)} x_j \tag{1}$$

where $0 < \beta < 1$ represents the spillover effect. The above definition does not specify the key factors to form the link, and it can be explained by factors such as the firm's location in space, R&D alliances, and R.J.V.[7] We will explain this phenomenon in detail in the implications section. Moreover, the above formula indicates that the R&D spillover can theoretically propagate throughout the entire network (after attenuation). As mentioned in the introduction, this outcome reflects an indirect spillover effect with the development of information technology and financial tools. However, the information will be distorted during the propagation process, and the R&D approach may be different for firms, so adjusting the spillover factor is necessary.

The payoff is defined as follows:

$$\pi_i = R_i - C_i = f\left(x_i + \sum_{l=1}^{n-1} \beta^{l-1} \sum_{j \in V_i^l(g)} x_j\right) - c \cdot x_i - k \cdot \left|V_i(g)\right| \tag{2}$$

where $f(\cdot)$ is a strictly concave incremental function, which indicates that the marginal revenue of R&D investment decreases. Furthermore, we assume that $f'(0) > c + 1$, which ensures that firms have strictly positive R&D investment, and there will be no situation in which firms do not perform R&D. The unit cost of R&D investment is c, including, for example, wages and equipment. The unit cost of establishing these links is k. The unit cost k will also change with the network $V_1(g)$. is a set of firms directly linked to firm I in a graph with adjacency matrix

[7] For example, firms can establish links through innovation alliances, supply chains, and spatial agglomeration. As the existing literature repeatedly demonstrates that spatial distance is related to technology spillovers, technology spillovers based on geographic distance can be expressed as.

Suppose that there is a certain threshold at work. If the geographical distance between firms is less than this threshold, it is assumed that there is a connection between firms; if it is greater than this threshold, it is assumed that there is no connection.

g (note that this graph is a directed graph). For example, in Fig. 1–1, $V_1(g) = \{2\}$ and $V_4(g) = \{\varnothing\}$.

3.2 Discussion About the Assumption

In this section, we discuss several key assumptions in the model setup and their theoretical foundation. The first issue concerns technology investment and technology spillover. In most studies about R&D games, technology spillover is substituted for the firm's R&D investment, such as D'Aspremont & Jacquemin (1988) and other studies following its initial setting. As a result, we use this setting so that the R&D input and network behavior (to obtain technology spillover) can be substituted with each other. However, taking a perspective from information systems (I.S.) research, there is a close link between spending on R&D and attaining spillover such that innovation capabilities can increase by gaining synergies between these two types of activities/investments. To respond to this concern, we extend the model to a multiple-stage game where the absorptive capacity of technology spillover will be strengthened by the accumulation of R&D investment in the previous stage. The model can be found in Appendix C. Although the network structure does not change, R&D investment gives the firm a higher incentive to engage in cooperative R&D and attain technology spillover.

The second issue concerns the diffusion effect. In the model, we assume that the technology spillover will spread to the whole network with attenuation β. We give two aspects of this background in the introduction: the technology information and the financial tools. Technology and finance are not necessarily increasing firms' ability to imitate or innovate but provide a particular environment. In such an environment, all firms can obtain technology information, but R&D investment and the absorption of technology spillover still depend on strategic behaviors.

3.3 Equilibrium

Let $y = argmax\, f(y) - cy$ and x_i^* be the investment in R&D in equilibrium. When the cost of establishing them is too high, i.e., k is large, firms will not establish links with one another. This is the first kind of equilibrium:

Equilibrium1: Suppose that $c \cdot y < k$; the equilibrium is as follows: $x_i^* = y$ for all and for all $i \in V$ and $g_{ij} = 0$ for all$i,\ j\ \in V$, where $y = arg\ max\, f(y) - cy$

This equilibrium is the most straightforward equilibrium, which means that all firms invest in R&D and no technology spillover in the network. When k decreases, it becomes profitable to establish links and absorb technology spillovers, which is the second type of equilibrium the rest of the paper focuses on.

Equilibrium2: Suppose that ; the equilibrium is one of the following:

2.1 (1).$\sum x_i^* = y$;(2)$\overline{g}_{i_1 i_2} = 1,\ g_{ji} = 1,\ g_{ji} = 0,\ g_{j_1 j_2} = 0\ for\ all\ i,\ i_1, i_2 \in\ I\ and\ j,\ j_1,\ j_2\ \in\ J$ where $I = \left\{ i\ \in V \middle| x_i^* > 0 \},\ J = \left\{ j \in V | x_j^* = 0 \right\} \middle| \right\}$.

2.2 (1)$x_i^* = \frac{k}{(1-\beta)c}$; (2) $g_{ij} = 0,\ g_{j_1 j_2} = 0,\ max\ d = 2\ for\ all\ i, i_1, i_2\ \in I\ and\ j, j_1, j_2\ \in J$

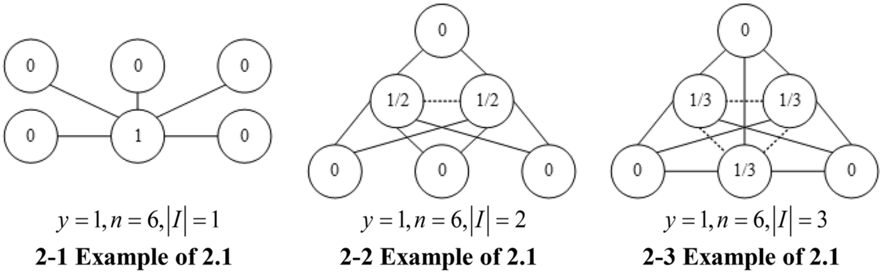


Fig. 2 Figs. 2-1, 2-2 and 2-3 example of 2.1. **Note:** The solid line in the figure represents the links established between type j and those of type I, and the dotted line represents the links established between type i. The numbers in the circles represent the R&D investment

where $I = \left\{ i \in V \middle| x_i^* > 0\},\ J = \left\{ j \in V | x_j^* = 0 \right\} \middle| \right\}$

The equilibria seem complicated while describing a simple structure that the distance between any two firms should be less than 2, where equilibrium 2.1 is the distance of 1 and equilibrium is the distance of 2. To meet other conditions of equilibria such as individual rationality, the R&D investment should be adjusted, leading to the difference of x_i^* in equilibrium 2.1 and 2.2.

The intuition of Equilibrium 2.1 is that all firms with positive R&D investment will establish links with one another and that all firms with zero R&D investment will establish links with firms with positive R&D investment, meaning that they form a "central-peripheral" structure. Figure 2 depicts several examples of this equilibrium. Note that there will be multiple R&D investment equilibria, but the firms will nevertheless form a similar network structure. We will then further analyze the number of firms that reside in the "center" of the network (i.e., the number of firms *i*) and the number of firms that reside in the "periphery" of the network (i.e., the number of firms *j*). The intuition of Equilibrium 2.2 is that all positive R&D investments are equal. Firms with zero R&D investment will establish links with some of the firms with positive R&D investment. Contrary to Equilibrium 2.1, the network structure in equilibrium is not unique, but the R&D investment is unique in equilibrium.

The equilibrium 2 seems to raise a paradox that the impact of technology spillovers is far-reaching while the distance of firms is short enough so that firms do not take advantage of the distant spillover. The explanation is that strategic behaviors limit other networks. If some firms obtain R&D spillovers from firms that are farther away, it is bound to mean that firms on the path can obtain more than the firms off of it. R&D spillovers reduce R&D investment incentives, which may lead to either insufficient investment in overall R&D by other firms or the firms will abandon the existing network structure and seek a lower-cost network structure for themselves.

Consider the following example. Firm A could absorb the spillover from a distant firm C. At this time, firm B in the path between firm A and firm C could absorb more spillover than firm A and firm C because the path of B-C and B-A is shorter than the path of A-C. Firm B would reduce its R&D investment and reduce the benefit from links between A and B, B and C. The distance between any two firms should be short enough to avoid this situation.

We prove the equilibria in the rest of this section. We will first prove two lemmas, which give some restrictions to the firms' strategies and reduce the variety of the equilibrium network. Later, we complete the proof with four parts. The first part proves the critical feature of the network that the distance between any two firms should be less than 2. The second part proves the conditions that the R&D investment should satisfy under the network. The third and fourth parts use the conditions mentioned above to prove equilibrium 2.1 and equilibrium 2.2.

Lemma 1 (Galeotti and Goya 2010): $x_i^* + y_i^* = y$ *for all* $i \in V$, where $y^* = \sum \beta^{l-1} \sum x_j^*$ represents the spillover absorbed by firm i.

Proof of Lemma1:

Suppose that. We have $f'\left[x_i^* + \left(\sum \beta^{l-1} \sum x_j^*\right)\right] > c$, which means that the firm can increase its payoff by reducing x_j^* is equilibrium.

Similarly, suppose that $x_j^* + y_i^* < y$. We have $f'\left[x_i^* + \left(\sum \beta^{l-1} \sum x_j^*\right)\right] > c$, which means that the firm can increase its payoff by increasing x_j^*. This contradicts that is equilibrium. Thus, for all firms:

$$x_i^* + y_i^* = y \tag{3}$$

3.4 Q. E. D

Lemma 2 Let $L = \{1, 2, \ldots, l\}$ be the vertex set in a loop. There is at least one true statement as follows (we set i+1=1 if i=l and if i=1):

(i) $\exists i, i' \in L$ such that $g_{i,i'} = 1$ and $x_j \geq x_{i'}, j = \frac{V_i(\overline{g})}{\{i\}}$

(ii) $\exists i, \in L$ is even and $\exists j, \in L$ is odd such that $g_{j,j+1} = 1,\ g_{j,j-1} = 1\ \textit{and}\ x_j < x_{j+1},\ x_j < x_{j-1}$

(iii) $\exists i, \in L$ is even and $\exists j, \in L$ is odd such that $g_{j,j+1} = 1,\ g_{j,j-1} = 1\ \textit{and}\ x_j < x_{j+1},\ x_j < x_{j-1}$

Proof of Lemma2:

The proof is divided into three parts.

Part1: we will prove that if the cardinality of the set L is odd, statement (i) is true.

Suppose that this statement is not true; we can assume $g_{1,2} = 1$ without generality and $x_2 > x_3$.

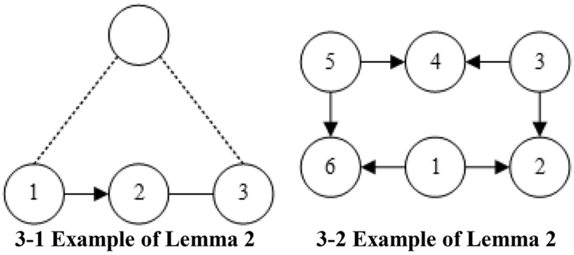
Then, we must have $g_{3,4} = 1$ and $x_4 > x_5$ since $x_2 > x_3$. Through this iteration, we have $x_{2i} > x_{2i=1}$. Consider vertex l. We have $x_{l-1} > x_1$ and $g_{l,1} = 1$. Through similar iterations, we have $x_1 > x_2,\ g_{2,3} = 1$ and finally $x_{2i-1} > x_{2i}$. Therefore, we have the following inequality:

$$x_1 > x_2 > \cdots > x_{l-1} > x_1 > x_1 \tag{4}$$

We have a contradiction.

Part 2: we will prove that if the cardinality of the set L is even and $g_{1,2} = 1$, statement (i) or (ii) is true.

Fig. 3 Figs. 3-1 and 3-2 example of Lemma 2



Suppose that statement (i) is not true. From the first part of the proof, we know that $x_{2i} > x_{2i+1}$. If $x_1 > x_2$, we have $x_{2i} > x_{2i-1}$ and the same contradiction as in (4). Therefore, we have $x_1 > x_2$ and $g_{1,l} = 1$. Through iteration, we have and. By combining and, we have and, which means that statement (ii) is true.

Part 3: we will prove that if the cardinality of the set L is even and $g_{2,1} = 1$, statement (i) or (iii) is true.

Through the following function $f : L \to L'$:

$$f\begin{cases} i+1, & i \text{ is odd} \\ i-1, & \text{is even} \end{cases} \tag{5}$$

We can map L onto L'. Statement (i) or (ii) is valid for L' from the second part of the proof. Thus, statement (i) or (iii) is valid for L.

3.5 Q.E. D

Statement (ii) and statement (iii) essentially reflect the same situation. Lemma 2 shows that in a loop, we must have a case like Fig. 3–1 or Fig. 3–2. In Fig. 3–1, there is at least one node (node 1) pointing to another (node 2), and the neighboring node (node 3) of node 2 has an R&D investment greater than or equal to that of node 2. In Fig. 3–2, all nodes with odd numbers point to adjacent nodes with an even number, and the R&D investment of even nodes is strictly more remarkable than that of the odd nodes. Next, we will begin to prove Equilibrium 2.

Proof of Equilibrium2:

The proof is divided into four parts.

Part1: We will prove that $\{d(i_1, i_2 \in I)\}$ for all $i_1, i_2 \in I$, where $d(i_1, i_2, \overline{g})$ is the distance (the length of the shortest path) between i_1 and i_2 in $\overline{g}$. This means that any two firms in the network can be linked through no more than one firm.

Suppose that $max\{dd(i_1, i_2, \overline{g})\} = 1 > 2$, and let $i_1 = 1$, $i_2 = l$, which means there is a path between firm 1 and firm l. Let $L = \{2, 3, \ldots, l-1\}$ be the set of firms on this path, where the distance between the firm and (firm 1, firm l) is . By lemma 1, for firm 1:

$$x_1 + x_2 + \beta^{l-1}x_1 + y_1' = y \tag{6}$$

where y_2' is the R&D investment firm 1 absorbs from firms except for firm 2 and firm l. Similarly, for firm 2, we have:

$$x_1 + x_2 + \beta^{l-2} x_1 + y_2' = y \tag{7}$$

where y_2' is the R&D investment firm 2 absorbs from firms except for firm 1 and firm l. By combining this with (6) and (7), we have $y_1' > y_2'$. Note that any R&D investment that firm 1 absorbs can be obtained by firm 2 (after adjusted by the spillover effect). Thus, there must be a firm r such that the distance between it and firm 1 is r ($d(1, r) = r$) and the distance between it and firm 2 is more significant than r ($d(2, r) > r$). Let $R = \{2, \ldots, r-1\}$ be the set of paths connecting firm 1 and firm r. The distance between firm 2 and the firm in L is less than the distance between firm 1 and the firm in L. Therefore, we have $L \cap R = \emptyset$.

Then, we consider two cases separately. In the first case, some firm m in R connects to some firm t in L. Then, the distance between firm 2 and firm t plus the distance between firm t and firm m must be greater than r:

$$t - 2 + 1 + r - m > r \tag{8}$$

Transposing the terms, we have:

$$t \geq m + 2 \tag{9}$$

Now, the distance between firm 1 and firm l through the path is:

$$l - t + 1 + m \leq l - 1 \tag{10}$$

This contradicts $d\left(1,\ l,\overline{g}\right)\} = l$ because $d\left(1, l, \overline{g}\right)\}$ is the shortest path between firm 1 and firm l.

In the second case, there is no firm in R connecting to firms in L, which means that $\overline{g}_{rt}$ for all $m \in R,\ t \in L$. Then, there will be four subcases.

Subcase 1: Firm 1 is on the path between firm r and firm l. The firm set on the path is L by definition, and we have$d(l, r) = r + l > l$. This contradicts $max\left\{d\left(i_1, i_2, \overline{g}\right)\right\} = 1$.

Subcase 2: Firm 1 is not on the path between firm r and firm l, and the distance between them is less than l-r($d(l, r) = r + l > l$). Then, the distance between firm 1 and firm l through firm r is $d(1, l) = d(l, r) + d(r, l) < l$, which contradicts $d(1, l = 1)$.

Subcase 3: Firm 1 is not on the path between firm r and firm l, and the distance between them is more than l-r($d(r, l) > l - r + 1$). Then, the distance between firm (the first firm on the path) and firm l is$d\left(R_2, l\right) = d\left(R_2, r\right) + d(r, l) > r - 1 + l - r + 1 = l$, which contradicts $max\left\{d\left(i_1, i_2, \overline{g}\right)\right\} = 1$.

Subcase 4: From the subcases above, we know the distance between firm r and firms l is either l-r or l-r+1. There is no connection between firms in R and firms in L. Firm 1, firm l, firm and firm form a loop. To simplify the proof, we let$P = L \cup \{1\} \cup \{l\} \cup \{r\}$ be the set of firms in the loop. Let firm 1 be the "start" of the loop and re-index the firms in counter-clockwise order, where the cardinality of P is odd or even, representing $d(r, l) = l - r$ or $d(r, l) = l - r + 1$, respectively.

By lemma2, if $|P|$ is odd and more prominent than 5, then there exists $g_{i,i'} = 1$ such that $x_j \geq x_i, j = \frac{V_i(\overline{g})}{\{i\}}$. Assume that g_{12}=1 and $x_3 \geq x_2$ without loss generality. For firm 1, the R&D investment it absorbs is:

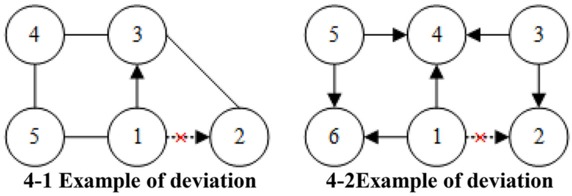


Fig. 4 Figure 4-1 Example of deviation in the figure, firm 1 can cut its link with firm 2 and establish a link with firm 3. Figure 4-2 Example of deviation in the figure, firm 1 can cut its link with firm 2 and establish a link with firm 4

$$x_2 + \beta x_3 + \cdots + \beta^{\frac{(p-1)}{2-2}} x_{\frac{(p-1)}{2}} + \cdots + \beta x_{p-1} + x_p \tag{11}$$

If firm 1 cuts the link between it and firm 2 and establishes a link with firm 3, the cost does not change, but the R&D spillover becomes:

$$\beta x_2 + x_3 + \beta x_4 + \cdots + \beta^{\frac{(p-1)}{2-1}} x_{\frac{(p-1)}{2}} + \cdots + \beta x_{p-1} + x_p \tag{12}$$

(12) is strictly larger than (11), so this kind of network is not the equilibrium. The intuition is that when firm 1 establishes a link with firm 3 instead of firm 2, this shortens the distance from any other firm in the loop. In addition, we have $x_3 \geq x_2$. The deviation gives a strict positive payoff to firm 1, which means that if |*P*| is odd, |*P*| must be less than 5 (the deviation is shown in Fig. 4–1).

Again, by lemma 2, if |*P*| is even and statement (i) is true, firm 1 can still increase its payoff by changing the network in equilibrium. Thus, statement (i) must be false, and statement (ii) or (iii) must be true. Because (ii) and (iii) describe the same situation, we assume that statement (ii) is true.

Note that $x_i = x_{i'}$ for all firm $i, i' \in P$ with an even index. Otherwise, firm x_{i-1} can cut its link with firm *i* and establish a link with firm *i'* to increase the technology spillover. If $|P| \geq 6$, firm 1 can deviate for a higher payoff through a similar strategy to that mentioned above. Specifically, firm 1 can play a strategy similar to Fig. 4–2. It cuts the link with firm 2 and establishes a link with firm 4. Now the R&D investment that firm 1 absorbs from firm 2, firm 3 and firm 4 does not change. In addition, the technology spillover from firm 5 becomes, which is strictly greater than before. It indicates that if is even, must be less than 6.

In conclusion, |*P*| is 3 or 4 and $max\{d(i_1, i_2, \overline{g})\}$ is 1 or 2 in the loop. This completes the first part of the proof.

Part 2: we will prove that if $d(i_1, i_2, \overline{g}) = 2$ all $x_i^* > 0$ are equal.

From the first part of the proof, we know that |*P*|=1 if . From lemma 2, suppose that not all are equal and the network in Fig. 5–1 is the only possible network structure with x_2>x_1,x_3 and x4>x_1,x_3 (otherwise, firm 1 can cut its existing link and establish one with firm 4). In the network, the R&D investment absorbed by firm 2 is:

$$y_2' = x_1 + x_3 + \beta x_4 \tag{13}$$

The investment absorbed by firm 1 and firm 3 is:

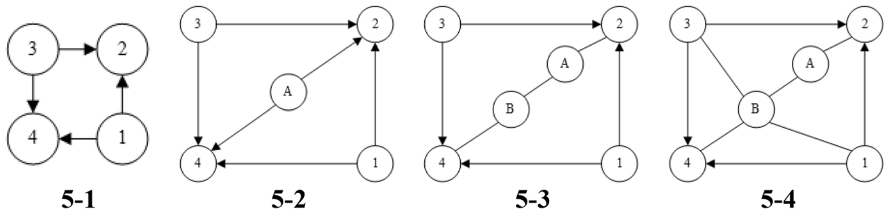


Fig. 5 Figs. 5-1, 5-2, 5-3, 5-4 example of First and Second Proof

$$\begin{aligned} y_1' &= x_2 + x_4 + \beta x_3 \\ y_3' &= x_2 + x_4 + \beta x_1 \end{aligned} \tag{14}$$

y_1' and y_3' are strictly larger than y_2', which means there must be some firms outside the loop. The distance between firm 2 and them is 1, and the distance between firm 1 and firm 3 and them is 2, which increases the spillover absorbed by firm 2. Assume that A is one of the firms. Since $max\{d(i_1, i_2, \overline{g})\} = 2$, firm A can either establish a link with firm 4 (as in Fig. 5–2) or establish a link with some other firm, such as firm B, and firm B links with firm 4 (as in Fig. 5–3 or in Fig. 5–4).

In Fig. 5–2, firm 2, firm3, firm 4, and firm A form a loop. The spillovers absorbed by the firms are not equal, as shown in (13) and (14). Therefore, there must be some other firms. The distance between firm 2 and them is 1, and the distance between firm 1 and firm 3 and them is 2. These firms form a new loop, and there is a similar situation in the new loop. This means that there is an infinite number of firms in set V, which is a contradiction.

In Fig. 5–3, firm 2, firm 3, firm 4, firm A and firm B form a loop. As shown in the first part of the proof, this kind of network is not an equilibrium. In Fig. 5–4, firm 2, firm 3, firm A and firm B form a loop, where the spillovers are not equal, as in Fig. 5–2. We thus obtain a similar contraction. This completes the second part of the proof.

Part 3: The first part of the proof states that $d(i_1, i_2, \overline{g}) = 1$ or $d(i_1, i_2, \overline{g}) = 2$. Part 3 of the proof focuses on the case in which, which is Equilibrium 2.1.

Suppose that $\sum x_i^* < y$. By $d(i_1, i_2, \overline{g}) = 1$, we have:

$$x_i^* + y_i^* = \sum x_i^* < y \tag{15}$$

This contradicts lemma 1. Now suppose that $\sum x_i^* < y$. We have:

$$x_i^* + y_i^* = \sum x_i^* < y \tag{16}$$

This also contradicts lemma 1. For the specific links with the firms, we have $\overline{g}_{i_1 i_2 = 1}$ directly from $d(i_1, i_2, \overline{g}) = 1$. Hence, the payoff is negative for $i \in I$ establishing links with $j \in J$. Therefore, we have g_{ij}=0.

For g_{j1j1}=1, we suppose that g_{j1j1}=0. Then, we have $d(j_1, i_1, \overline{g}) = 1 > 1$ and:

$$\sum_{i \neq i_1} x_i^* + \beta^{l-1} x_{i_1}^* < \sum_{i \in I} x_i^* = y \tag{17}$$

This contradicts lemma 1. Therefore, $g_{ij} = 1$ for all $j \in J$ and $i \in I$. It also means that the payoff is negative for $j_1 j_2 \in J$ to establish links with one another and $g_{j_1 j_2} = 0$. This completes the third part of the proof.

Part 4: we will prove that $d(i_1, i_2, \overline{g}) = 2$ i.e., Equilibrium 2.2.

We can directly obtain $g_{ij} = 0$; otherwise, firm I can cut its link with firm j and link with some firm $i' \in I$ to absorb more R&D investment. Similarly, we have $g_{ij} = 0$. Furthermore, by lemma 1, the R&D investment absorbed by firm j is exactly y in equilibrium, which is equal to the R&D investment absorbed by firm i. Hence, firm j must establish links with firm *i's* neighbor with positive R&D investment. The distance between any two firms is less than 2.

Now, we turn to the R&D investment in equilibrium. Figure 6 is a subgraph in the equilibrium where firms 1 ~ 4 have positive R&D investments. We have proven that if firm j has a link with firm 1, it also has links with firm 1's neighbors, i.e., firm 2 and firm 4, which means that the payoff for firm j in this network must exceed that in any other network. Assume that firm j cuts its link with firm 2 (as in Fig. 6–1) or cuts its links with firm 2 and firm 4 and forms a link with firm 3. In both cases, the cost decreases by k, and the technology spillover also decreases by$(1 - \beta)x_i^*$. Therefore, we have the following:

$$k \leq (1 - \beta)cx \tag{18}$$

Otherwise, firm j will cut its links and invest in R&D itself. Moreover, firm 1 does not establish a link with firm 3 (as in Fig. 6–3), which means that the cost of establishing a link (k) exceeds the cost of investing in R&D:

$$k \geq (1 - \beta)cx \tag{19}$$

By combining (18) and (19), we have:

$$x = \frac{k}{(1\beta)c} \tag{20}$$

This completes the fourth part of the proof.

3.6 Q.E.D

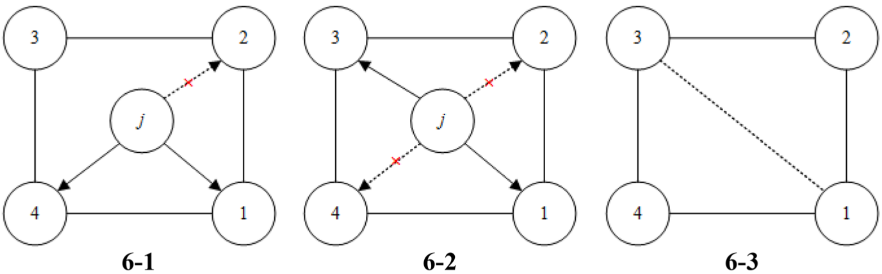


Fig. 6 Figs. 6-1, 6-2 and 6-3 example of Fourth Proof

4 Endogenous Innovation Networks and Implications

In the baseline model, we abstractly define variables such as firm networks and R&D investments. This section will accord specific economic meaning to the above variables and extend the equilibrium results to derive their implications. We consider three different forms of corporate networks. The first is based on the network formed in cooperative R&D, where the connections and technology spillovers between firms depend on disseminating knowledge and information. The second represents spatial agglomeration and innovation parks (Silicon Valley, for instance). In this case, the links between firms and technology spillovers are established by geographical distance. Finally, we focus on firms' information capital and analyze whether the network will be centralized or decentralized under a specific innovation policy.

We pay special attention to the endogenous network structure and its influence. Network structure can be described from many different perspectives. Jackson et al. (2017) classified network structure from macro and micro perspectives. The macroscopic characteristics of the network include the scale and density of the network. The first part is about cooperative R&D, and the second part is about spatial aggregation, which mainly focuses on these characteristics of the network structure and their influence. The micro characteristics of a network include the location and centrality of nodes or groups. The third part focuses on the analysis of information capital, mainly focusing on a network's characteristics.

Of course, network structure also includes many other attributes, such as solid connection and weak connection, Gil Schmidt power centrality, and other different centrality (Mhera et al. 2006; Sinclair et al. 2009). We did not choose these characteristics to be included in the analysis, mainly because they are more common in the analysis of individual-based social networks, and it is not easy to give them an intuitive economic explanation in the context of innovation alliance. Therefore, we choose the scale, density, and other more intuitive features to draw more applicable conclusions.

4.1 Cooperative R&D networks

We define $i \in I$ as innovators due to their positive R&D investment and define $j \in J$ as imitators due to their zero R&D investment. Innovators share knowledge through cooperative R&D, such as R.J.V. In the process, the development of information technology enables firms to communicate more effectively, thus reducing the cost of establishing links between firms (Carson et al. 2003). Moreover, information technology development also enables imitators to obtain technology spillovers at a lower cost, generating new challenges for intellectual property protection (McManis 1996). We will use the previous equilibrium to compare the number of innovators ($|I|$) and the number of imitators ($|J|$) when the cost of the links changes. We analyze whether the more efficient technology spillovers generated by the development of information technology encourages innovation or imitation.

We first analyze the number of firms of the two types. The following Proposition 1 shows that the number of firms in I cannot be too large; otherwise, the spillover cost $k \cdot \left|V_i(g)\right|$ between firms will increase rapidly, which exceeds the research and development cost savings brought about by the collaboration. In Equilibrium 2.1, the links are established between any firms with positive R&D investment. Therefore, as the number of innovators increases, the spillover cost increases quadratically ($O\left(n_1^2\right)$), while the R&D cost savings decrease linearly ($O\left(n_1\right)$). The spillover costs will eventually exceed R&D cost savings. This conclusion is closely related to the network structure of firms in equilibrium. This network structure limits the R&D spillovers that firms enjoy from others farther away and increases the number of links that firms need to establish.

Proposition 1: *In Equilibrium 2.1* $n_1 \equiv |I| \le cy/k, n_2 \equiv |J| \ge (nk - cy)/k,$ *where* $I = \left\{i \in V|x_i^* > 0\right\}$ *and* $J = \left\{j \in V|x_j^* = 0\right\}$

Proof of Proposition 1: In Equilibrium 2, the number of links for firm $i \in I$ is:

$$\sum_{a=1}^{n_1-1} a = \frac{n_1\left(n_1 - 1\right)}{2} \tag{21}$$

Moreover, $cy_I^* \ge k\left|V_i\right|$ for all firm $i \in I$, and thus, we have:

$$\sum cy_i^* \ge k \sum \left|V_1\right| = \frac{kn_1\left(n_1 - 1\right)}{2} \tag{22}$$

By lemma 1, we have $x_i^* + y_i^* = y$. Hence, $\sum y_i^* = \left(n_1 - 1\right) \cdot y$ and plugging this it into (22) yields:

$$\left(n_1 - 1\right)cy \ge \frac{kn_1\left(n_1\right) - 1}{2} \tag{23}$$

By simplifying (23), we have:

$$n_1 \le \frac{2cy}{k}, n_2 = n - n_1 \ge \frac{nk - 2cy}{k} \tag{24}$$

For firm $j \in J$, the number of links is n_1. In addition to $cy \ge kn_1$ by the definition of equilibrium,[8] we have:

$$n_1 \le \frac{2cy}{k}, n_2 = n - n_1 \ge \frac{nk - cy}{k} \tag{25}$$

By combining (24) and (25), we complete the proof of Proposition 1.

4.2 Q.E. D

In Equilibrium 2.2, the numbers of innovators and imitators are also limited. The upper bound of the number of innovators is related to the spillover effect β and increases as the spillover effect decreases. This is mainly because, in Equilibrium

8 As we show in the proof of Equilibrium 2, if, firm can deviate to cut its links and invest in R&D itself.

2.2, the R&D investment of each firm is particular and directly related to the spillover effect β. When β is small, more firms will participate in cooperative R&D, increasing the number of innovators' upper bound. However, compared with Proposition 1, Proposition 2 also indicates that the number of innovators has an upper bound and a lower bound. This difference is also related to the network structure of Equilibrium 2.2. The R&D spillover absorbed by a firm is not limited to its neighbors, which requires at least a certain number of firms with a distance of 2 from the firm. When β decreases, the number of these firms should increase, and thus, the number of innovators has a lower bound.

Proposition 2: *In Equilibrium 2.2,* $\frac{(1-\beta)cy}{k} < n_1 \frac{(\beta^{-1}-1)cy}{k}, \frac{nk-(1-\beta)cy}{k} < n_2 < \frac{nk-(\beta^{-1}-1)cy}{k}$ *where* $n_1 \equiv |I|,\ I = \left\{i \in V | x_i^* > 0\right\}$ *and* $n_2 \equiv |J|,\ J = \left\{j \in V | x_j^* = 0\right\}$.

Proof of Proposition 2: For $i \in I$, let $I_i - 1$ be the number of a firm's neighbors with positive R&D investment and be the number of its neighbors' neighbors (except for itself) with positive R&D investment. From Equilibrium 2.2, we have:

$$\left(I_i + I_2\beta\right)x_i^* = y \tag{26}$$

Plugging in $x_i^* = \left[(1-\beta)cy\right]/k$ into it, we have:

$$I > I_1 + I_2\beta = \frac{(1-\beta)cy}{k}\beta I < I_1 + I_2\beta = \frac{(1-\beta)cy}{k} \tag{27}$$

By combining and simplifying the inequalities in (27), we complete the proof of Proposition 2.

4.3 Q. E. D

From Proposition 1 and Proposition 2, it can be found that in Equilibrium 2, the number of innovators decreases as the cost of spillovers k increases, and the number of imitators increases as k increases. The opposite is true for the R&D cost c. This seems to contradict economic intuition: it is generally believed that imitators need to establish links with other firms, and as the cost of spillovers increases, the number of imitators should be reduced. Proposition 1 and Proposition 2 show that the above logic applies to imitators and innovators. As the cost of spillovers increases, innovators will find that the benefits of cooperative R&D are decreasing and that independent R&D is more profitable, leading imitators to connect with fewer innovators and reduce the cost of imitation.

Furthermore, the conclusion responds to the empirical research on network structure and innovation performance to a certain extent. Capello and Faggian (2005) and Giuliani (2013) found that the agglomeration of enterprises is conducive to knowledge spillover and dissemination; in other words, a denser network is conducive to the transmission, diffusion, and absorption of information, thus improving the innovation performance of firms. Rowley et al. (2000), Gilsing et al. (2008) document that intensive linkages may damage firms' innovation performance or have an inverted U-shaped relationship. In the more intensive network, information may tend to be homogeneous, and new knowledge and information are more difficult to penetrate the innovation alliance, thus hindering innovation, especially breakthrough innovation.

This conclusion reveals that enterprises' innovation network is formed by enterprises based on information dissemination and diffusion conditions (such as technical conditions). With the increase of the convenience of information diffusion, innovators will form a network structure that is more conducive to disseminating information (the number of innovators increases and contacts increase) and reducing the innovation activities of each innovator, which implies a new influence mechanism or path of information diffusion to network structure and innovation performance.

Regarding policy practice, this analysis challenges the following "common sense" to some extent: more convenience in information dissemination makes more firms become imitators, and intellectual property protection becomes more difficult. However, the proposition shows that the propagation conditions' convenience reduces the imitator's cost and reduces the innovators' cost of cooperative R&D, and the group engaged in collaborative innovation becomes larger. For example, when firms can easily share knowledge, a firm can undertake only a small part of the work in a large research project, which increases the imitative difficulty for the imitators and facilitates intellectual property protection. Innovation policy supports collaborative innovation with firms but has a negative attitude towards horizontal mergers for R&D reasons. With the new communication tools, the standards and definitions of "copy" and "imitation" change concerning intellectual property rights protection. We argue that these standards and definitions may be less strict because the spread of technology information also helps protect to some extent.

4.4 Spatial Agglomeration of Innovation Networks

A technology park or agglomeration, such as Silicon Valley, brings together many innovative firms. Silicon Valley's success has also led to its imitation in India and China, and similar technology clusters have been established. The spatial attenuation of technology spillovers is an important reason why R&D firms form a spatial agglomeration (Desmet and Rossi-Hansberg 2014). Traditionally, one considers a firm as the center, and technology spillovers will benefit firms within a specific geographical distance, but beyond this distance, technology spillovers will stop. Other reasons for the spatial agglomeration include the sharing of complementary resources and personal relationships among entrepreneurs. For example, the initial establishment of *Fairchild Co.* in Silicon Valley produced an advanced semiconductor, an essential resource for many high-tech products at that time, such as microprocessors, televisions, and cameras. As a result, many firms formed an R&D cluster in the Bay Area, ultimately creating the Silicon Valley area. However, technologies such as cloud computing and SaaS make it less necessary for firms to agglomerate in a particular area and allow them to obtain resources in distant areas, e.g., firms can use the computing capacity of Amazon by accessing the Internet. Does it mean that there is also no need for firms to pursue spatial agglomeration?

We first provide the conclusion, followed by the analysis. The cluster becomes less integrated when the geographic spillover of technology is less attenuated. Meanwhile, the cluster remains integrated when the distance is fixed, but technology spillover can spread from firm to firm. The former exists in the high-tech industry, where research resources such as data are virtual and can be shared through the Internet. The latter exists in traditional industries that implement an information system (as discussed in the introduction), where research resources such as materials are physical, and the distance still matters.

Consider the first case, in which the spillover distance increases. Assume that the payoff becomes

$$\pi_i = R_i - C_i = f\left(x_i + \sum_{l=1}^{n-1} \beta^{l-d} \sum_{j \in V_i^l(g)} x_j\right) - c \cdot x_i - k \cdot \left|v_i(g)\right| \tag{28}$$

where the spillover effect of a single connection β^{l-d_i} is different for firm i. A larger d_i indicates that technology spillover will decrease less with further distance. In other words, firms in a distant area will absorb more technology spillover. By Lemma 1, we can directly derive Corollary 1:

Corollary 1: Firms will connect with other firms only if $c \cdot \beta^{l-d_i} > k$.

If there exists d so that all firms $d_i > d$ and Corollary 1 holds, the corollary is the same as lemma 1, and firms will form a single connected graph implying geography integration. In contrast, when d_i is small, firms will reduce connections with others. Therefore, the strengthening of spillover d_i (by distance) increases the spatial agglomeration.

Next, we consider the second case, where the technology spillover can spread from firm to firm. Based on Equilibrium 2.1 and Equilibrium 2.2, we find that with the strengthening of spillover, a new clustering model (Equilibrium 2.2) formats, and the network is less centralized. Therefore, the strengthening of spillover will not increase the spatial agglomeration.

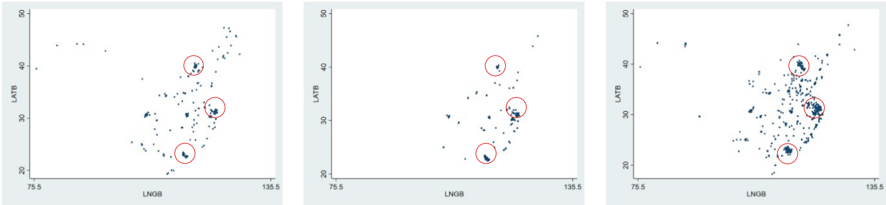
Finally, we provide some empirical evidence of the previous analysis. We use the data on new establishments in China in 2017, including the agriculture, textile, and e-commerce industries. Agriculture is an industry that is less influenced by technology spillover changes because production and innovation are highly correlated with land use. We use it as a baseline. E-commerce is the industry where the distance of technology spillover expands. Firms can share resources through the Internet and break the distance limits; this scenario corresponds to the section's first case. Textiles are industries that are influenced mainly by the information system. The spillover distance remains fixed because the physical material should be transported among cooperative firms and is still impacted by the distance; this scenario corresponds to the second case. The data processing and detailed analysis methods are presented in Appendix D.

Fig A ~ Fig C present the location of establishments in different industries. There are three main industry clusters in China: the Beijing (the top circle in the figure), Yangtze River (the middle circle), and Zhu River (the bottom circle) clusters. In the

Table 1 Proportion of the number of firms in clusters

	Beijing	Yangtze River	Zhu River	Total
Agriculture	14.71%	12.65%	6.76%	34.12%
Textile	8.12%	25.37%	25.62%	59.11%
E-commerce	12.20%	23.76%	10.53%	46.46%

figure, we find that agriculture firms' locations are dispersed. In textiles, firms are more integrated into the three main clusters. However, in e-commerce, firms gather in the three clusters, while many firms are in other areas.



A: Agriculture B: Textile C: E-commerce

Besides, we calculate the proportion of firms in the three clusters. The results are presented in Table 1 and are similar to those in the figure. We find that there are a large number of firms in the clusters in textiles (59.11%), while the proportions are smaller in agriculture (34.12%) and e-commerce (46.46%). This finding supports the findings in this section that firms in textiles are still integrated while firms in e-commerce are dispersed.

We summarize the part of theoretical findings and empirical evidence in Table 2. In E-commerce, the spillover strengthens when firms can share knowledge in the far distance, and the theoretical model finds the network is less integrated. In comparison, the spillover strengthens in the textile industry when firms can share knowledge with more upstreams or downstreams using information systems. The model predicts that the network is highly integrated. The spatial agglomeration of establishments in the industries supports the findings.

4.5 Industry Structure of Innovation Networks

Information capital refers to firms' ability to access or disseminate information through network connections (Jackson 2018). In R&D, having higher information capital means that a firm occupies a relatively "central" position in the network, and it can obtain information from other firms. The information capital here differs from the information technology capability of the firm. It measures not the

Table 2 Information Diffusion and Clusters

	Technology Spillover	Network	Integration in data
Agriculture	Baseline (not influenced)	Disperse	34.12%
E-commerce	Spillover strengthen by distance	Less integrated	46.46%
Textile	Spillover strengthen by firms	High integrated	59.11%

R&D equipment and R&D personnel owned by the firm but the specific network structure in which the firm is located. If the information capital gap between firms is large, the degree of network centralization is high. Here, we analyze how changes in parameters such as R&D costs and connection costs can affect the distribution of information capital in the network.

We first calculate the degree of firms in equilibrium, and then we introduce the method for measuring information capital. We take the following symbolic representation:

Let I_1 and I_2 be the set of firms with $x_i^* > 0$ in Equilibrium 2.1 and Equilibrium 2.2, respectively. We also use them to represent the cardinal of the set when there is no confusion. (2) Let $I_2(1) \subset I$ and $I_2(2) \subset I$ be the set of firms with a distance of 1 and 2 from firm *I*, respectively. We also use them to represent the cardinal. (3) Let $D(i) = \left|V_i(\overline{g})\right|$ be the degree of the firm I in the undirected graph and $D^1(i) = \left|V_i^l(\overline{g})\right|$ be the number of firms with the distance of l from the firm I in the undirected graph.

*Proposition 3: In Equilibrium*2.1,$D(i) = N - 1$ *for* $i \in I_1$ *and* $D(j) = I_1$ *for* $j \in J_1$.

In Equilibrium 2.2,$I_2(1) = \frac{y - I_2 \beta x_i^*}{(1-\beta)x_i^*}, I_2(2) = \frac{I_2 x_i^* - y}{(1-\beta)x_i^*}$, $D(i) = I_2(1) + \gamma$, $D^2(i) = I_2(2), D^3(i) = J_2 - \gamma$ *for* $i \in I_2$ *where* γ *is a constant with a range of* $0 \le \gamma \le J_2$

$D(j) = I_2(1) + 1, D^2(j) + \varphi, D^3(j) = J_2 - \varphi$ *for* $j \in J_2$ *where* φ *is a constant with a range of* $0 \le \varphi \le J_2 - 1$

Proof of Proposition 3: In Equilibrium 2.1, firms I_1 in link with one another. Firms J_1 in link with firms I_1 in but have no links with one another. Therefore, we have that the degree of firm I is $D(i) = I_1 - 1 + J_1 = N - 1, i \in I_1$ and the degree of firm j is $D(j) = I_1$.

In Equilibrium 2.2, all firms in invest equally in R&D. By lemma 1, for any firm $i \in I_2$:

$$I_2(1)x_i^* + \beta\left[I_2 - I_2(1)\right]x_i^* = y \tag{29}$$

We solve the following equation:

$$I_2(1) = \frac{y - I_2\beta x_i^*}{(1-\beta)x_i^*} I_2(2) = I_2 - \frac{y - I_2\beta x_i^*}{(1-\beta)x_i^*} = \frac{I_2 x_i^* - y}{(1-\beta)x_i^*} \tag{30}$$

For firm I in J_2, its neighbors include firms in and firms in directly linked to it. Therefore $D(i) = I_2(1) + \gamma$, we have, where γ is a constant representing the firms J_2 in directly linked to firm i. The firms with a distance of 2 from firm I are $I_2(2)$, so we have $D^2(i) = I_2(2)$. The firms with a distance of 3 from firm I include $I_2(2)$ and firms with no direct links with it in J_2. Therefore, we have $D^3(i) = J_2 - \gamma$.

For firm j in, as we show in Equilibrium 2.2, it has a link with an arbitrary firm I and its neighbors to satisfy the condition that the technology spillover is exactly y, so the neighbors of firm j include firm I and all its neighbors, and the degree is $D(j) = I_2(1) + 1$. The firms with a distance of 2 from firm j are $I_2(2)$ and firms in J_2, which also have the same links with firms in I_2 as firm j. Therefore, we have$D^2(j) = I_2(2) + \varphi$, where φ is a constant. The firms with a distance of 3 from firm j are the rest of the firms in J_2, and hence $D^3(j) = J_2 - \varphi$.

4.6 Q.E.D

Next, we introduce the definition of information capital. Following Jackson (2010), we define information capital as follows:

$$Dec_1\left(\overline{g},\ p\right) = \sum p^1 D^1(i) \tag{31}$$

where p represents the loss of information in transit, which may be due to distortions in the transmission of information or the exchange of information between two firms with a probability less than 1. Note that this definition is similar to the definition of technology spillovers in (1). In fact, technology spillovers can also be understood as the transfer of technology information between firms, and the information will gradually become distorted through the spillover, where β represents the degree of information distortion. We obtain Proposition 4 directly from Proposition 3:

Proposition 4: In Equilibrium 2.1, the information capital of firm I is $Dec_i = D(i) = N - 1$ and the information capital of firm j is $Dec_j = D(j) = I_1$.

In Equilibrium 2.2, the information capital of firm I is

$Dec_i = D(i) + pD^2(i) + p^2D^3(i) = \frac{y}{x_i^*} + p^2J_2 + \left(1 - p^2\right)\gamma, i \in I_2$

and the information capital of firm j is

$Dec_j = D(j) + pD^2(j) + p^2D^3(j) = \frac{y}{x_i^*} + p\left(1 - p^2\right)\varphi + p^2J_2 + 1, j \in J_2$

Combining the above with Proposition 1 and Proposition 2, we find that with a decrease in the spillover cost k and an increase in the R&D cost c, $|I|$ will increase and $|J|$ will decrease. In Equilibrium 2.1, the information capital of firm $i \in I$ is unchanged, and the information capital of firm $j \in J$ is reduced, meaning a more centralized network. In Equilibrium 2.2, the information capital of firm $i \in I$ and $j \in J$ that of firm are simultaneously reduced. From the perspective of the firm, its location is less characterized by "centrality." By extending this logic to all firms, we find that the network is decentralized. Hence, the impact of cost changes on centralization will be bipolar because the initial network structure is different. In networks with a high initial degree of centralization (Equilibrium 2.1), cost changes with a decrease in k and increases in c will further enhance the degree of centralization, which means that a small number of firms are more prominent in cooperative R&D. In the multicluster network structure with weak initial centralization (Equilibrium 2.2), the cost change will further weaken the degree of centralization, and more firms will participate in cooperative R&D.

We first discuss the role and influence of network centrality. Centrality measures the degree of network or the distribution of information capital. If information capital is concentrated in a few enterprises, the network's centrality is more substantial. Existing studies have drawn different conclusions on the relationship between network centrality and performance. Grund (2012) found that a decentralized network is more conducive to cooperation among members; that is, there is a negative relationship between centrality and performance. On the contrary, Tsai and Ghoshal (1998) believe that centrality promotes the perceived trust between enterprises, conducive to exchanging information and innovation between enterprises.

The improvement of leader centrality in the network can improve other members' performance in the network significantly (Mhera et al. 2006).

The centrality is the cause of the degree of cooperation and information diffusion, and its result is the specific network structure formed by the network members to adapt to the external environment. In the multicluster network, the increase of information diffusion (measured by parameter k) makes members rely more on cooperation than independent innovation. However, due to multiple clusters in the network, the centrality of innovator clusters does not improve, making the network centrality decline, and the innovation activities of each enterprise will also decrease. In the center-periphery network, the increase of information diffusion makes more enterprises become innovators rather than imitators, connecting more around innovators and bringing about network centrality. However, the increase of innovators will also reduce the innovation activities of each enterprise at this time.

In the face of the change of external environment, the centrality change of the multicluster network is different from that of the center peripheral network, which also brings about the change of the network position and the innovator and imitator's power groups. From the perspective of the regulation of innovation alliance or merger and acquisition, previous research and practice paid more attention to the positive role of the alliance in integrating information dissemination but more or less ignored the impact of different network structures on market power.

The information capital of a firm involved in research and development is also a source of market power. The above analysis reveals how the entire industry's exogenous impact affects the R&D behavior of a firm and thus affects market power. In studies of innovation policy, there is no consistent conclusion regarding whether cooperative R&D is beneficial or detrimental to social welfare (Leahy et al. 1997; Miyagiwa 1997). The model in this paper argues that depending on the initial network structures present in different industries, innovation policy may have diametrically opposite effects. One of the reasons for cooperative R&D, such as joint research ventures, is that alliances will save the cost of communication (k in the model) or save the cost of R&D through the sharing of resources (c in the model). This section shows that cooperation will make fewer firms take prominent positions and potentially increase market power if the network is initially a core type. As a result, such alliances should be carefully examined or have certain limitations. However, if the network is initially a multicluster type, cooperation will make a more significant number of firms participate in R&D and will potentially increase welfare, providing a reason for cooperation.

5 Conclusions and Discussion

5.1 Conclusions

We analyze the network structure formed by endogenous technology spillovers in an R&D game model. We further apply the equilibrium results to innovators and imitators in collaborative R&D, the spatial agglomeration of firms, information capital, and market power. We find that the connections between firms are incredibly close; that is, the distance between two firms is no more than two. The first network structure forms a center for innovative firms, while the rest of the firms revolve around these firms and absorb their R&D spillovers. The second network structure forms multiple centers for innovative firms, and the rest choose a center to link to and absorb R&D spillovers.

Regarding the spatial agglomeration of firms, with the development of information technology, technology spillovers are no longer limited to a specific geographical distance but can spread to distant places. However, the equilibrium network limits firms' technology spillovers. Firms still tend to form spatial agglomerations. However, they may not gather around a single center at this time but may form multicentre clusters.

Regarding information capital and market power, the firms' locations in a network will have different impacts on other firms, which is also a market power source. In a more centralized network, a small number of firms will have high market power. The opposite is true in networks with lower levels of centralization. We find that depending on the initial network structure and the degree of centralization, a change in research and development costs and technology spillover costs, among other factors, will have the diametrically opposite effect on the degree of centralization. This conclusion implies that innovation policy needs to consider the initial network structure of cooperative R&D networks. Depending on the initial network structure, innovation policy or industry regulation may have opposing effects in different industries.

5.2 Discussions

Previous theoretical and empirical studies have found that network structure affects information transmission or knowledge sharing in innovation alliance, affecting innovation performance. This paper's model reveals another influence path: the network structure is formed to adapt to the specific environment. In particular, even after improving information transmission conditions, firms' strategic innovation behavior to avoid "free-riding" will make the innovation alliance still maintain a high network density and centrality. The endogenous adjustment of strategic innovation behavior and network structure limits information diffusion. This endogeneity also means that when considering the impact of patent protection policy and M&A review, we should consider that the network structure may also be adjusted and affect the policy's effect in practice. Furthermore, our model suggests that endogenous networks simplify the networks and extract the key features of the innovation structures in practice.

This study also distinguishes the "innovator" group and "imitator" group in innovation alliance, which are different in innovation behavior, network location, and characteristics. In many previous studies related to innovation alliance, the internal innovation alliance is homogeneous or homogeneous (although there are differences in the network structure between Innovation Alliances), so that the behavior or performance of each individual in the innovation alliance can also represent the behavior and performance of the whole innovation alliance (Gilsing et al. 2008). This study shows that the internal innovation alliance is not uniform; alliance members occupy different positions and play different roles. Some recent empirical studies support this argument from the side (Ryu et al. 2018): enterprises outside innovation alliance may have more contact with competitors, and innovation alliance will reduce this impact in various ways. Besides, with the increase of information diffusion, the innovation alliance forms a multicluster network, and the imitators may occupy a more central network position than the innovators. For future research, systematic research on the behavior, function, location, and performance of heterogeneous groups within alliances can become one direction.

This paper mainly considers the interaction between the network diffusion effect (technology spillover) and network structure. Future research can further analyze the interaction between more factors and network structure. For example, when the uncertainty is introduced, it would be interesting to investigate how the network structure can be adjusted to better deal with the risk dispersion and what is the position of the central enterprise when multiple innovation alliances are competing under different degrees of technology spillovers, and how the network structure balance the agglomeration effect of R&D and the potential leakage risk.

Acknowledgements We benefited from comments and suggestions of seminar participants at the Renmin University of China, Nanjing Audit University, Zhejiang University, University of California, Berkeley, University of San Diego, and participants in various workshops and conferences. Lihong Yang acknowledges financial support from the National Natural Science Foundation of China (NSFC: 71573269), the Major Project Funding for social science research in Jiangsu Province (Grant No. 2021SJZDA123), and the Research Institute of Alibaba Ant Group. Shixun Wang thanks Major Project Funding for social science research base in Fujian province social science planning, the Social Science Foundation of Fujian Province in China (Grant No. FJ2020B118). This article started while authors were visitors at the University of San Diego and the Hass School of Business at the University of California, Berkeley, the hospitality of whom we gratefully acknowledge. All errors are our own.

Appendix A: Standard two-stage R&D model

In the standard two-stage R&D model, firms decide to invest in R&D in stage 1 and compete in the product market in stage 2. In appendix A, we will prove that the baseline model can extend to the two-stage R&D model.

The inverse demand function is.

R&D investment will reduce the unit cost in production:

where is the revenue function in the basic model. The profit of the firm is.

The problem in the first stage is

The problem in the second stage is

By backward induction, the solution to the second-stage problem is.

Plug this solution into the first-stage problem, and.

where, and are parameters consisting of. When there is no technology spillover, the optimal R&D investment is.

Let where y is the optimal R&D investment.

When technology spillover exists, there is a substitution between its R&D investment and the technology spillover absorbed from other firms. Hence, Lemma 1 still holds, and the total R&D input is. Otherwise, the firm will reduce its R&D

investment when the total R&D input exceeds and increase its R&D investment when the total input is less than. Lemma 2 also holds, as does equilibrium 2.

Appendix B: Innovation Spillover and Uncertainty

In Appendix B, we consider a situation in which the firm's innovation is not certain and is affected by a random variable. This random variable can represent both the uncertainty of the R&D process and the heterogeneity among firms. Many studies note that the uncertainty of R&D makes R&D spillovers occur mainly in innovations (Miyagiwa and Ohno 2002; Erkal and Piccinin 2010). R&D investment itself does not directly generate spillover effects, but spillover can be achieved only after conversion to innovation. In appendix B, we will prove that the baseline model can be extended to the symmetric equilibrium in uncertainty.

The innovation of the firm is defined as follows:

where is the R&D investment of firm *i* and is a random variable drawn from some c.d.f. The spillover that firm i absorbs from firm j is.

We set if in continence. The payoff of the firm is:

where represents the firms with maximum innovation in the connecting subgraph, the nodes of which contain firm *i*. The technology spillover is the innovation from other firms. If the innovation from spillover is more advanced than its own innovation, the firm will absorb this innovation. Otherwise, the spillover has no impact (Desmet and Rossi-Hansberg 2014). The expectation of the spillover is.

Let, and the lemma can be expressed as in equilibrium. If, the reduction in will decrease and increase and thus increase payoff. If, the increase in will increase and decrease. The firm can increase until even if.

Lemma 2 still holds, and the main idea in proving equilibrium 2 also holds. Suppose the maximum distance between any two firms is larger than 2 in network. There must be a loop to satisfy. By lemma 2, the number of nodes in the loop cannot exceed 4, which is a contradiction. Therefore, the maximum distance between the firms is 1 or 2.

When the maximum distance is 1, any firms with positive R&D investment will establish links with each other, according to modified lemma 1 above. The links with firms with zero R&D investment have no impact on, so there is no link with firm *j*. This is the case for equilibrium 2.1. When the maximum distance is 2, the technology spillover absorbed by firms should also be equal and is equal in equilibrium, which is the case of equilibrium 2.2.

Appendix C: Dynamic Model of the Spillover Effect

We consider a multiple-stage game. In the first stage, firms decide on R&D investment and a network link to maximize their current payoff. In the following stage, the R&D investment in the previous stage will benefit the technology spillover. We use the model to capture the close link between R&D investment and technology spillover.

To obtain economic intuition, we first consider the game in stage 2. The action remains the same as that in stage 1, and the payoff of firm *i* is.

where is an increasing function and is the R&D investment in the previous stage.[9] Therefore, a higher R&D investment in stage 1 will make the technology spillover in stage 2 more effective.

Changing the cost of attaining technology spillover will change whether to join an innovation alliance and absorb spillover but will not change the decision of which firm to connect with. To see this phenomenon, the cost to absorb technology spillover is different from. Therefore, if, firm *i* will solely invest in R&D and will not connect with other firms. Firm *i* with saves costs for the firm to connect with others and absorb technology spillover. The cost of connection is identical among all firms. Therefore, it will not change which firm's decision to connect with compared to the decision in stage 1. From this analysis, we derive Corollary C1.

Corollary C1: Firms will connect with other firms in stage t only if , where , and are the corresponding parameters in stage t.

The corollary builds a link between historical behavior and the current network structure. For example, consider the parameters following a function of time:, and. After a particular stage, the firms can be imitators and absorb technology spillover only if innovators are in the previous stage. More findings can be derived from different parameter settings, such as subjecting each parameter to a specific distribution.

The link between R&D investment and technology spillover in this study has several implications. First, the R&D investment previously gives the firm a higher incentive to engage in cooperative R&D and attain technology spillover. Second, the two types of networks in the basic model (core and multicluster) do not change, given that all other parameters remain the same.

Appendix D: Empirical Evidence of Spatial Agglomeration

We use the data on new establishments in China in 2017, including the agriculture, textile, and e-commerce industries. Based on the establishment address information, we use XGeocoding and Baidu Map's API to obtain each firm's latitude and longitude. We use the longitude and latitude as the x-axis and y-axis, respectively, to draw the scatter figure in Fig A ~ Fig C.

From the figure, we can find that the firms gather mainly in three areas, China's main clusters. They are clusters of Beijing (the top circle in the figure), Yangtze River (the middle circle), and Zhu River (the bottom circle). Thus, we further calculate how many firms are in the clusters. We present the results in Fig A ~ Fig C in Sect. 4.2.

[9] To simplify the notation, we cancel out the time subscript.

Due to data limitation, it is difficult to obtain the exact boundary data of longitude and latitude in each cluster, so we directly use the address information to identify whether the firm belongs to a specific cluster. The cluster of Beijing mainly covers the areas in Beijing. The cluster of the Yangtze River covers mainly the areas in Shanghai, Zhejiang, and Jiangsu provinces. The Zhu River cluster mainly covers the areas in Guangdong province. Therefore, if the address contains the string "Beijing," we consider the firm to belong to the Beijing cluster. A similar process is performed for the Yangtze River and Zhu River clusters.

Then we aggregate the number of firms in each cluster and divide it by the industry's total number of firms. We present the results in Table 1 in Sect. 4.2.

References

Acemoglu D, Akcigit U, Kerr W (2016) Innovation Network. Proc Natl Acad Sci 113(41):11483–11488

Aggarwal VA (2020) Resource congestion in alliance networks: how a firm's partners' partners influence the benefits of collaboration. Strateg Manag J 41(4):627–655

Amir R, Evstigneev I, Wooders J (2003) Noncooperative versus cooperative R&D with endogenous spillover rates. Games Eco Behav 42(2):183–207

Aoyagi M (2018) Bertrand competition under network externalities. J Eco Theory 178:517–550

Borgatti SP, Halgin DS (2011) On Network Theory. Organization Science 22(5):1121–1367. https://doi.org/10.1287/orsc.1100.0641

Bramoullé Y, Kranton R, D'amours M (2014) Strategic interaction and networks. Am Eco Rev 104(3):898–930

Calvó-Armengol A, Patacchini E, Zenou Y (2009) Peer effects and social networks in education. Rev Econ Stud 76(4):1239–1267

Capello R, Faggian A (2005) Collective learning and relational capital in local innovation processes. Reg Stud 39(1):75–87

Carson SJ, Madhok A, Varman R, John G (2003) Information processing moderators of the effectiveness of trust-based governance in interfirm R&D collaboration. Organ Sci 14(1):45–56

Cassiman B, Veugelers R (2002) R&D cooperation and spillovers: some empirical evidence from Belgium. Am Eco Rev 92(4):1169–1184

Chu Y, Tian X, Wang W (2018) Corporate innovation along the supply chain. Mngt Sci

d'Aspremont C, Jacquemin A (1988) Cooperative and noncooperative R & D in duopoly with spillovers. Am Econ Rev 78(5):1133–1137

Desmet K, Rossi-Hansberg E (2014) Spatial development. Am Econ Rev 104(4):1211–1243. https://doi.org/10.1257/aer.104.4.1211

Elliott M, Golub B, Jackson MO (2014) Financial networks and contagion. Am Eco Rev 104(10):3115–3153

Erkal N, Piccinin D (2010) Cooperative R&D under uncertainty with free entry. Int J Ind Organ 28(1):74–85. https://doi.org/10.1016/j.ijindorg.2009.07.003

Federico G, Langus G, Valletti T (2018) Horizontal Mergers and Product Innovation. Int J Ind Organ 59:1–23

Gagnon J, Goyal S (2017) Networks, markets, and inequality. Am Eco Rev 107(1):1–30

Galeotti A, Goyal S (2010) The law of the few. Am Eco Rev 100(4):1468–1492

Gil Moltó MJ, Georgantzis N, Orts V (2005) Cooperative R&D with endogenous technology differentiation. J Eco & Mngt Strat 14(2):461–476

Gilsing V, Nooteboom B, Vanhaverbeke W, Duysters G, van den Oord A (2008) Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. Res Policy 37(10):1717–1731. https://doi.org/10.1016/j.respol.2008.08.010

Giuliani E (2013) Network dynamics in regional clusters: evidence from Chile. Res Policy 42(8):1406–1419

Grund TU (2012) Network structure and team performance: The case of English Premier League soccer teams. Social Networks 34(4):682–690

Hagedoorn J (2002) Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. Res Policy 31(4):477–492. https://doi.org/10.1016/S0048-7333(01)00120-2
Jackson MO (2010) Social and economic networks. Princeton university press
Jackson MO (2018). A typology of social capital and associated network measures
Jackson M, Rogers B, Zenou Y (2017) The Economic Consequences of Social-Network Structure. J Eco Literature 55:49–95
Jan, van, den, Ende, Geerten, & van, et al (2012) The paradox of standard flexibility: the effects of co-evolution between standard and interorganizational network. Organ Stud 33(5–6):705–736
Kamien MI, Muller E, Zang I (1992) Research joint ventures and R&D cartels. The Am Eco Rev, 1293–1306
Kamien MI, Zang I (2000) Meet me halfway: research joint ventures and absorptive capacity. Int J Ind Organ 18(7):995–1012
Kireyev A, Leonidov A (2018) Network Effects of International Shocks and Spillovers. Netw Spat Econ 18(4):805–836
Leahy D, Neary JP (1997) Public policy towards R&D in oligopolistic industries. The Am Econ Rev, 642–662
Lee J (2010) Heterogeneity, brokerage, and innovative performance: endogenous formation of collaborative inventor networks. Organ Sci 21(4):804–822
Li H, Tang Y (2019) Network Structure and Dynamics of Chinese Regional Incubation. Netw Spat Econ 19(4):1173–1197
Manea M (2018) Intermediation and resale in networks. J Polit Econ 126(3):1250–1301
McManis CR (1996) Taking TRIPS on the Information Superhighway: International Intellectual Property Protection and Emerging Computer Technology. Vill l Rev 41:207
Mehra A, Dixon A, Brass D, Robertson B (2006) The Social Network Ties of Group Leaders: Implications for Group Performance and Leader Reputation. Organ Sci 17(1):64–79
Miyagiwa K, Ohno Y (1997) Strategic R&D policy and appropriability. J Int Econ 42(1–2):125–148
Miyagiwa K, Ohno Y (2002) Uncertainty, spillovers, and cooperative R&D. Int J Ind Organ 20(6):855–876. https://doi.org/10.1016/S0167-7187(01)00079-0
Qiao T, Shan W, Zhang M, Liu C (2019) How to facilitate knowledge diffusion in complex networks: The roles of network structure, knowledge role distribution and selection rule. Int J Inf Manage 47:152–167
Roijakkers N, Hagedoorn J (2006) Inter-firm R&D partnering in pharmaceutical biotechnology since 1975: Trends, patterns, and networks. Res Policy 35(3):431–446. https://doi.org/10.1016/j.respol.2006.01.006
Rowley T, Behrens D, Krackhardt D (2000) Redundant governance structures: an analysis of structural and relational embeddedness in the steel and semiconductor industries. Strateg Manag J 21(3):369–386
Ryu W, Mccann BT, Reuer JJ (2018) Geographic co-location of partners and rivals: implications for the design of r&d alliances. Acad Manag J 61(3):945–965
Schilling MA, Steensma HK (2001) The use of modular organizational forms: analysis. Acad Manag J 44(6):11149–11168
Sinclair PA (2009) Network centralization with the Gil Schmidt power centrality index. Social Networks 31(3):214–219
Tsai W, Ghoshal S (1998) Social capital and value creation: The role of intrafirm networks. Acad Manage J 41(4):464–476. https://doi.org/10.5465/257085
Tsiotas D, Polyzos S (2018) The complexity in the study of spatial networks: an epistemological approach. Netw Spat Econ 18(1):1–32. https://doi.org/10.1007/s11067-017-9354-1
Vives Xavier (2009) Oligopoly pricing: old ideas and new tools. M.I.T. press
Xie X, Fang L, Zeng S (2016) Collaborative innovation network and knowledge transfer performance: A fsQCA approach. J Bus Res 69(11):5210–5215
Zhang Q, Zhang J, Zaccour G, Tang W (2018) Strategic technology licensing in a supply chain. Eur J Oper Res 267(1):162–175