Research Policy 50 (2021) 104117

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

A man is known by the company he keeps?: A structural relationship between backward citation and forward citation of patents

Wonchang Hur [a], Junbyoung Oh [b,*]

[a] Professor, College of Business Administration, Inha University, 253, Yonghyun-dong, Nam-gu, Incheon 402-751, Republic of Korea
[b] Professor, Department of Economics, Inha University, 253, Yonghyun-dong, Nam-gu, Incheon 402-751, Republic of Korea

ARTICLE INFO

JEL Classifications:
H23
O31
O38

Keywords:
Patent value
Social network analysis
Network structure of backward citation

ABSTRACT

Inventing is a recombinant process that involves searching and recombining different streams of knowledge. The value of invention is associated with not only how many prior inventions are considered, but also how they are related to each other. We introduce social network analysis broadly used in the social capital theory, and extend the dimension of analysis for the evaluation of patent value. This study employs U.S. pharmaceutical patent data and investigates whether the network characteristic of backward citations have significant effect on the future patent value. The empirical results suggest that the network features of backward citations measured by constraint, cohesion, and efficiency have statistically significant implication on the value of invention in both level and depreciation rate. The study also provides empirical evidence that the exploration strategy is more significantly and positively correlated with the future value of invention compared to the exploitation strategy of inventors.

> ``*In general, innovations are new combinations of existing knowledge. To produce other things, or the same things by a different method, means to combine materials and forces differently.*''
>
> J. Schumpeter, 1934, pp. 65–66

1. Introduction

The complexity of modern innovation increasingly necessitates the pooling and integration of multiple strands of knowledge. In addition, the ability to explore the potential of prior inventions and combining them with new ideas is critical for a firm's innovation capability (Ahuja and Lampert, 2001; Subramaniam and Youndt, 2005; Cohen and Levinthal, 1990; Iansiti and West, 1997; Christensen, 2002).[1] In this sense, innovation is like a continuous process of ``search and recombination,'' and the value of created technologies is dependent on both the characteristics of the search scope and the inventor's own capability (Arthur, 2007; Savino et al., 2017).

Since the seminal work of Jaffe et al. (1993) on technology spillover, many scholars have used patent citations as a primary instrument to assess the value of an invention and its economic impact. (Caballero and Jaffee, 1993; Jaffe et al., 1993; Trajtenberg et al., 1997; Harhoff et al., 1999; Lanjouw and Schakerman, 2001; Hall et al., 2005; Mehta et al., 2010; Jung and Lee, 2016; Moser et al., 2017; Corredoira and Banerjee, 2015; Kuhn et al., 2020). Most existing models, however, have relied on the individual citation information listed on patent document, and relatively less attention has been paid to the interrelationship among citations, which serves as a valuable source of information for understanding how a new invention is created from previous inventions (Carpenter et al., 1981; Lerner, 1994; Tong and Frame, 1994; Lanjouw and Schankerman, 2001; Allison and Lemley, 2000; Harhoff et al., 2003; Sampat et al., 2003).

This study employs the social capital theory in which the ego's performance is determined by not only his or her own characteristics but also their relationships with alters. The social capital perspective highlights that ego's relationships with alters should be considered as a part of the larger relationship network formed by the interactions between alters and other alters (Burt, 2004; Ahuja, 2000; Reagan and Zuckerman, 2001; Rosenkopf and Nerkar, 2001; Vasudeva et al., 2013). We take this social capital perspective as our theoretical framework to

* Corresponding author.
E-mail addresses: wchur@inha.ac.kr (W. Hur), jun@inha.ac.kr (J. Oh).

[1] In the literature on organizational learning and strategic management, such ability has been termed 'combinative capability' (Kogut and Zander, 1992), 'dynamic capability'' (Teece, Pisano, and Shuen, 1997), and 'architectural competence' (Henderson and Cockburn, 1994)

https://doi.org/10.1016/j.respol.2020.104117
Received 28 March 2018; Received in revised form 31 August 2020; Accepted 31 August 2020
Available online 6 September 2020

understand whether and how the structure of prior art (which is identified by the citation network among cited inventions) is related to the future value of an invention. We particularly focus on the critical tradeoff between recombining a cohesive set of prior knowledge and a heterogeneous set of prior knowledge. A patent founded on a cohesive set of prior knowledge may produce redundant information and, therefore, lacks novelty with a high risk of replacement by many similar alternatives and will have only limited impact on subsequent patents. On the contrary, such homophily of prior knowledge may have merits in the technology spillover perspective, because relying on a cohesive body of knowledge can relieve the burden of heterogeneity and promote the reliability of the created knowledge. The empirical evidence in the literature on these contrasting views is rare, and, to the best of our knowledge, we introduce for the first time a new dimension of analysis on this issue with solid empirical evidence.

For empirical analysis, we employ U.S. pharmaceutical patent data granted during 2001~2005 that includes citation information. By referring to network measures from social network analysis, we introduce the three network variables of *constraint, cohesion, and efficiency* that can effectively characterize the structural patterns hidden in the citation relationships among prior inventions and examine how they are related to patent value. The variables are calculated from each patent's ego network of backward citations generated by both the 1st order- and the 2nd order-backward citations (i.e., citations among cited documents). Following the previous literature, which suggests that applicant citation is more relevant to the knowledge flow invention, and examiner citation is more relevant to the private value of the invention, we separate applicant citations from examiner citations in measuring network variables (Alcacer and Gittelman, 2006; Hedge and Sampat, 2009).

The value of an invention is defined from two perspectives: private value for the owner, and social impact or technological spillover on subsequent inventions. These aspects are measured by forward citations and patent renewal. For forward citations, we consider two variations, simple count and survival probability, to examine the level and the depreciation of patent value. For estimation, we employ the negative-binomial, Cox-Hazard ratio, log-normal survival probability, and logistic regressions corresponding to the dependent variables used.

The empirical results show that the structural feature of the prior art has a statistically significant correlation with a patent value. Three network variables, *constraint, cohesion,* and *efficiency,* have statistically significant implications for forward citations in both the level and the depreciation rate. The estimated effects are substantial. For example, with all other things being equal, a 1% increase in *constraint* tends to decrease the expected count of examiner forward citations by 7.20% (7.89% for total forward citations) and increase the deprecation rate of examiner forward citations by 4.85% (6.91% for total forward citations). The network structure of backward citations also shows a significant correlation with the probability of patent renewal. A 1% increase in *constraint* decreases the odds of renewal by 2.58%. These results show that the network structure of backward citations has important information in predicting the future value of a patent, and its absence leads to a serious overestimation or underestimation.

The remainder of this paper is organized as follows. In Section 2, we briefly review previous studies on social capital theory and network measure of the prior art. In Section 3, we introduce the methodologies employed in this empirical study and discuss econometric models. Section 4 provides the empirical results, and Section 5 concludes the study.

2. Literature review

The main goal of the study is to characterize the network structure of a patent's prior art and to examine whether the prior art structure has implications for the predictability of its technological impact or future value. This question is motivated by the structural social capital studies that normally use a network approach, in which the structural relationship among actors is a key element in understanding their behavior and performance.[2] In this section, we first review two competing perspectives in the structural social capital literature and discuss their implications in understanding the relationship between the prior art and the inventive process.

2.1. The network closure perspective vs. structural hole perspective

Structural social capital is one of three dimensions of social capital, along with cognitive and relational social capital (Nahapiet and Ghoshal, 1998). Structural social capital is the network of people whom an individual knows and upon whom he or she can draw for benefits such as information and collaboration. Just as other types of capital (e.g., human, physical, financial), structural social capital is perceived as an asset that can be used for value creation. It creates value by facilitating an individual's accessibility to external parties for acquiring knowledge and by increasing opportunities for people to gain access to relevant peers with desired sets of knowledge or expertise. Therefore, structural social capital is embedded in the structure of a network, which is built from actors placed at certain positions in the network and interactions among them.

In the structural social capital literature, two competing views have been proposed: the 'network closure' and the 'structural hole' perspectives. A network with closure refers to a social network structure in which everyone is connected such that no one can escape the notice of others, which in operational terms, usually means a dense network (Coleman, 1988; Burt, 2004). Studies advocating the network closure perspective claim that social capital is created by a dense network of strongly interconnected members. Frequent interactions among network members make it easier to produce similar views and common beliefs, which has profound effects on group norms (Wang et al., 2010). Strong ties in networks with closure are, therefore, more likely to foster homophily within a group, which can promote reliance, collaboration, and knowledge sharing among its members.

The network closure perspective has analogous implications for searching and recombining technological knowledge in the invention process. A dense network of a patent's citations represents focus and coherence in a set of prior knowledge and could be evidence suggesting that the inventor appreciates the benefits of the network closure. In fact, inventors referring to a cohesive body of knowledge are effective in concentrating their learning efforts and absorbing deeper technological knowledge. Hence, the closed network structure among cited inventions may imply that inventors efficiently lower the technological risk, improve the communication of homogeneous technologies, and decrease the costs of adopting analogous technology.

The network closure in patent-citation networks is also related to the 'exploitation' hypothesis, which postulates that learning is cumulative, and its associated performance is maximized when the object of learning is related to what is already known (Cohen and Levinthal, 1990). In the context of invention, inventors following the exploitation strategy must be interested in deepening their knowledge in specific fields in which they are competitive. Therefore, the structure of prior knowledge they focus on is more likely to take the closure form. Several studies support this perspective. Halfat (1994) argues that firms tend to concentrate their R&D efforts in areas related to preexisting knowledge bases and

[2] Social network analysis has also been applied often to nonhuman relationships, such as communication and alliance of organizations and enterprises (Uzzi, 1997), generation and diffusion of innovating knowledge (Burt, 2004; Reagans and McEvily, 2003), and the evaluation of journal influence (Peng and Wang, 2013). Collins and Clark (2003) studied the dynamic relationship between human resource practice and firm performance in the context of social network relationships in a top management team and found that the effect of social network among team members is a moderate factor in firm performance.

tend to produce new knowledge closely related to existing knowledge in which they have current expertise. Silverman (1999) shows that a firm elects to enter markets in which it can exploit its existing technological resources and in which its existing resource base is strongest. Martin and Mitchell (1998) also point out that most firms, after having entered a product market, introduce similar designs to those of existing ones. Persistence in R&D is also related to the closure perspective. Jaffe's research into ``technological position'' (1986, 1989) found that firms benefit from ``nearby'' R&D far more than from ``distant'' R&D, suggesting limits on the fungibility of technological knowledge. Schoenmakers and Duysters (2010) also found that radical inventions are—to a higher degree—based on existing knowledge compared with non-radical inventions.

In contrast with the network closure perspective, the 'structural hole' perspective of social capital claims that the network closure may constrain the direction of new knowledge creation (Dosi, 1982; Teece et al., 1997). A structural hole is a missing or weak connection between two network members bridged by a broker. The structural hole perspective views that the broker position has an intervening opportunity to take advantage of the non-redundant relationships with each member, and thus, has more social capital. The position of the structural hole is a good place for mediation, forming a critical path connecting different actors and effectively controlling the flow of resources and information (Freeman, 1979). Burt (2004) argues that an actor located on the structural hole bridges the different information flows or allocates resources among distinct groups, creating a communication channel between actors who are not connected directly. In this brokering process, actors in the structural hole remain open to opportunities to create novelty. Relatedly, Henderson and Cockburn (1994) define 'architectural competence' as the ability to access new knowledge from outside the boundaries of the organization and the ability to integrate knowledge flexibility across disciplinary.

By analogy, structural holes in a network of cited inventions can provide inventors with similar opportunities as in the structural holes of social capital. For instance, a patent positioned at a structural hole in a network of cited inventions can be considered as an outcome of integrating knowledge from distinct fields. Such patents are likely to create value from the novel recombination of distinct technological fields. As Granovetter (1973) suggests in his theory on the strength of weak ties, inventors seeking knowledge from disconnected fields are likely to access a broader array of ideas and opportunities than those focused on a cohesive set of knowledge. Such boundary spanning strategy generates ``information benefits'' compared to staying within a cohesive set of knowledge in which redundant information is more likely to be generated.

Regarding the inventors' strategy for technological search, the structural hole perspective has significant implications. Inventors exploring a prior art knowledge network having many structural holes can move beyond a local boundary and feature many opportunities to reconfigure their knowledge base (Stuart and Podolny, 1996; Hargadon and Sutton, 1997; Danneels, 2002; Nesta and Saviotti, 2005). Such boundary-spanning exploration helps uncover useful technology options that have not been exploited. This enables inventors' trials of extensive technological options and enhances the possibility of finding novel combinations by taking advantage of cross-fertilization between different options (Granstand, 1998; Suzuki and Kodama, 2004). In this view, more technologically diversified inventions can capture more opportunities and technical possibilities (Nelson, 1959; Quintana-García and Benavides-Velasco, 2008).

Arguably, structural holes and network closure are not necessarily contradictory, but rather play different roles that are useful for different aims. Several efforts have been made to reconcile these contradictions. Kogut and Zander (1992) emphasize the balance of two countervailing strategies and argue that the process of search and recombination requires both exploiting a firm's own knowledge base and exploring relevant but unexplored areas. Levinthal and March (1993) highlight the trade-off between the two strategies based on the fact that the effectiveness of learning in the near neighborhood of current technologies interferes with learning that takes place at a distance. Inventors who engage exclusively in exploring external technology options will often suffer from the lack of returns on their own knowledge, while others who engage exclusively in exploitation will frequently suffer from obsolescence (Levinthal and March, 1993; Sorenson and Stuart, 2000). The exploitation strategy is more focused on coordination problems, while the exploration strategy focuses more on learning benefits from different skills, information, and experience (Coleman, 1988; Burt, 1992; Reagans and Zuckerman, 2001). This paper builds on this contingency view. Instead of taking one side of the two competing views, our primary objective is to examine how the structural characteristics of prior knowledge networks are associated with the outcomes of the invention process and the value of a patent.

2.2. *prior art and patent citations*

To apply for a patent, inventors are legally required to reveal relevant prior art available to the public for its claims of originality. In the United States, applicants (and their attorneys) have a ``duty of candor'' to disclose any prior knowledge, and if a person having a duty to disclose deliberately avoids reporting prior knowledge, then a patent can be found unenforceable due to inequitable conduct (USPTO, 1998, Section 2242). After applicants disclose their references, patent examiners conduct their own prior art searches under the patent granting procedure. If the examiner fails to uncover prior art material to reject patentability, applicants may receive a broader, stronger patent. This possibility gives an incentive for the applicants to selectively report prior art supporting their claims despite the ``duty of candor.''

Notably, the network structure of backward citations could have different implications depending on the types of citation used. Recent empirical studies have revealed that applicant citations have different characteristics from examiner citations (Alcacer and Gittelman, 2006; Hedge and Sampat, 2009; Alcácer et al., 2009; Cotropia et al., 2013). While applicants cite prior art that they recognize as being relevant to their claimed invention, examiners serve as ``devil's advocate'' in the patent prosecution process and cite prior art that an inventor missed in reporting (Lemley, 2001, p. 1502). Thus, examiner citations are more likely to block or limit the claimed inventions rather than reflect the knowledge spillover from the prior art. Alcácer and Gittelman (2006) find that examiners add a significant number of citations (40% of all citations), and their citations differ systematically from applicant citations. They argue that, in this case, the aggregate backward citations, which include examiner citations, may have significant 'noise' in measuring the knowledge flow that occurred to an inventor. Hedge and Sampat (2009) assess whether examiner- or applicant-forward citations accurately predict the private value of patents, which is measured by applicant payment of maintenance fees. They find that the applicant's prior art citations better represent the knowledge flow, while examiner citations are more closely related to the invention's private value.

On the other hand, Hedge and Sampat (2009) and Cotropia et al. (2013) speculate that applicants have incentives to selectively cite only patents that can support their inventions, while they are legally required to disclose any previous patents relevant to their application. If they knew of patents that clearly block their invention, they would likely either forgo patent protection or claim around these patents. Lampe (2012) also suggest that applicants tend to cite prior art that supports their claims, either because they deliberately withhold art that would invalidate their claims or because they are careful to draft their claims to avoid the prior art. Furthermore, inventors may not list all prior knowledge completely even if they don't have such misintention. The recent skyrocket increase in the number of registered patents makes it more difficult for inventors to identify all relevant knowledge for their invention, especially in the areas where technological convergence occurs frequently, and similar technologies have already been developed in other fields. In this case, applicants are likely to miss some part of the prior art even though much of the past knowledge is related to their invention. Jaffe et al. (2000) surveyed inventors regarding their knowledge of works cited in their patents and find that knowledge of works cited is often incomplete, and it is complemented by examiner citations during the patent examiner prosecution process. Jaffe et al. (2000)'s findings indicate that the applicant backward citation itself could be limited and may not enough to describe the characteristics of the knowledge base that a given invention stands on. Battkea et al. (2016) and Nemet and Johnsona (2012) also state that aggregate patent citations could be the best available indicator for knowledge flows because they provide accessible and comprehensive information about the linkages between patents, and the additions of examiner citation improve the picture of whence knowledge drives. Thus, the use of aggregate backward citations may compliment the potential incompleteness of inventor citations, particularly when the purpose of using backward citations is not limited to the identification of knowledge flows to inventors but is extended to identifying the technological position of an invention in a patent landscape of state-of-the-art technology. To address this issue, we conduct the supplementary analysis using the aggregate backward citations (including both applicant and examiner citations) to supplement the primary analysis using only application citations only.

3. Econometric strategy

3.1. Empirical model

The main goal of this study is to characterize the network structures of backward-cited patents of an invention and examine whether they are correlated with the *ex-post* patent value. Eq. (1) presents the regression equation used in this study. The dependent variable $Y_{ikt+\Delta t}$ indicates the *ex-post* value of patent *i* at year $t + \Delta t$, which was assigned to inventor *k* at year *t*. The time window of prediction, Δt, is a fixed constant for all *t* to address the truncation problem. **NW** represents the network variables created by backward citations and is of major interest in our study. The network features of backward citations are characterized by *constraint, cohesion*, and *efficiency*. ***X*** is the set of control variables and includes the number of backward citations, the number of Cooperative Patent Classification (CPC) codes in backward citations, the recency and the time spread of backward citations, and the number of claims. μ_k indicates the patent class dummy variable and θ_t indicates the year fixed effect dummies.

$$Y_{ikt+\Delta t} = \alpha + \beta \cdot NW_{ikt} + \gamma \sum_{l=1}^{m} X_{ikt} + \mu_k + \theta_t + \varepsilon_{ikt} \quad (1)$$

In addition to examining the effects on patent value, survival analysis is performed to examine the depreciation rate of patent value. The depreciation of patent value over time represents the obsolescence of the underlying technologies. A *Cox-hazard* model and a parametric survival model with a log-normal distribution are used for survival analysis. The survival analysis estimates the risk rate of citation termination at a specific time *t*. The risk ratio, which is called a hazard ratio, is modeled as follows:

$$\ln\left(\frac{H(t)}{H_0(t)}\right) = \sum X_i \cdot \beta_i \quad (2)$$

In Eq. (2), $H_0(t)$ is a baseline hazard when all independent variables are zero. The equation estimates how the network variables affect the hazard ratio of citation termination compared to the baseline hazard.

3.2. Measurements for patent value

A variety of patent value indices have been proposed in the literature, depending on the research contexts.[3] In general, the monetary value of a patent, if it could be quantified reasonably, would be the most straightforward measure since the ultimate goal of promoting inventions is to seek financial returns. However, the exact valuation of the expected financial returns is hardly a straightforward process. Commercializing an invention often involves a multi-stage process and is subject to highly uncertain market environments. Therefore, most approaches to valuation are subject to substantial bias in statistical analyses.

In this paper, the value of an invention is defined from two perspectives: private value for owners, and social impact or technological spillover on subsequent inventions. To measure the private value of a patent, we consider examiner forward citations and patent renewal. As mentioned prior, examiner forward citations are known to better reflect the private value of patents (Hedge and Sampat, 2009). Patent renewal data also provides direct information on the private value of protecting the proprietary rights, since the renewal data logs the payment records of periodic renewal fees that patentees must pay to keep their patents in force.

In the recombinant perspective on inventions, the social value of a patent consists of its spillover impact on subsequent inventions. An invention's spillover impact can be measured as the extent to which it affects subsequent inventions. A good proxy for quantifying the extent of spillover is the number of forward citations. In the innovation literature, forward citations have been widely accepted as a successful proxy for a patent's future technological impact (Albert et al., 1991; Shane, 2001; Dahlin and Behrens, 2005). Following this tradition, we measure $Y_{ikt+\Delta t}$ as *the total number of forward citations* patent *i* receives during the given period of time, Δt. Notice that we do not separate the examiner citations here because we consider the social value as encompassing the broader aspects of spillovers to both applicants and examiners. These social value measures effectively complement the use of examiner forward citation counts.

We also examine the *survival hazard ratio* of forward citations as additional dependent variables. The hazard ratio is interpreted as the chance of an event occurring compared to the baseline hazard. In our study, the survival hazard ratio indicates the risk that the forward citations discontinue in the following period. In order to supplement the proportional hazard estimation, we also estimate the survival probability by assuming that the survival time follows a log-normal probability distribution. While the proportional hazards model only provides a proportional ratio of the covariates' effects to the baseline hazard, a hazard model with a log-normal probability distribution can test whether covariates have accelerating or decelerating effects on the timing of the event of interest.

3.3. Measurement of prior art structure: 2nd order (alters-to-alters) citation networks

The structure of the prior art of a patent can be represented by its ego network of backward citations, which is the unit of our analysis. A patent's ego network consists not only of 1) its 1st order citations to previous inventions (that is, ego-to-alters citations) but also of 2) the 2nd

[3] In the literature, the patent value index is usually categorized into market-based indexes and patent-based indexes. The most well-known market-based indexes include Tobin's Q, market value of shares (enterprise level), royalties, inventors and managers' evaluation of patent and acquisition activity (patent level). In contrast, patent-based indexes are more diversified and include technological impact (forward citation), geographic importance (patent family), technological life (renewal), and legal dispute (litigation probability and opposition probability).

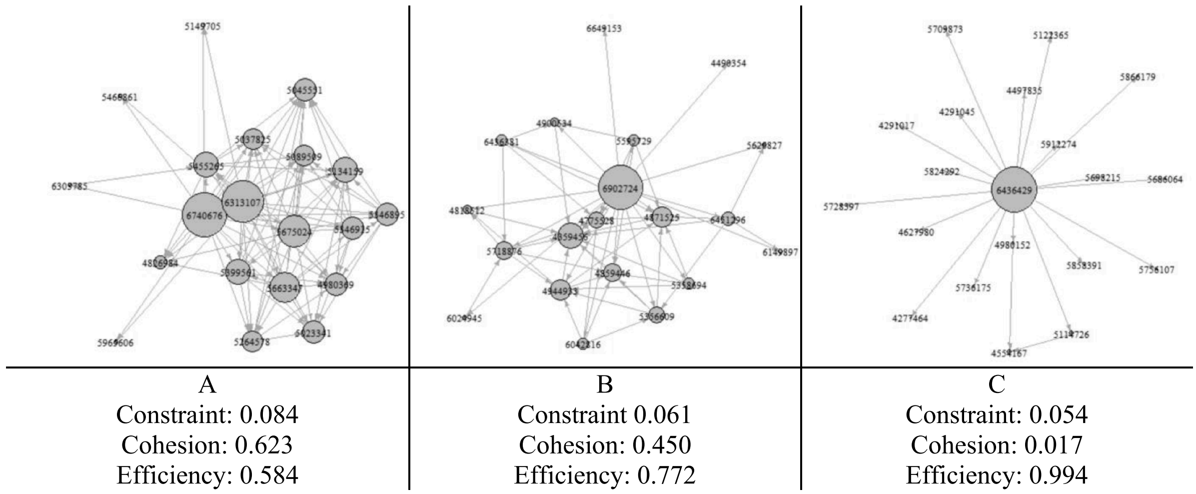
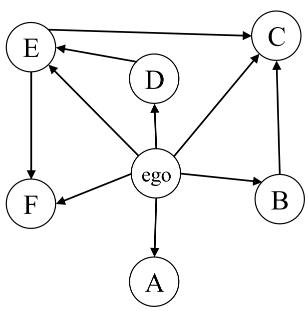


Fig. 1. Ego-centric Backward Citation Networks of Three Patents. Note: All patents have the same number (20) of backward citations, but the structure of citation network among alters themselves is more cohesive in patent A.



Alters (j)	Constraint $\left(P_{ij} + \sum_q P_{iq} P_{qj}\right)^2$	Efficiency 1 - $\sum_q P_{iq}\, m_{qj}$
A	0.09	6/6
B	0.22	5/6
C	0.36	4/6
D	0.24	5/6
E	0.44	3/6
F	0.24	5/6
Normalized sum	0.265 (1.59 / 6)	0.778 (4.67 / 6)

Fig. 2. An Example Calculation of Constraint and Efficiency. Note: 1) This figure illustrates how to calculate the constraint and the effective size of Ego. The direction of edge is from a citing patent to cited patents, which is the opposite to the direction of knowledge flow. P_{ij} denotes the proportion of i's relationship with j, which is 1 over the number of edges of i. m_{ij} is a binary variable denoting whether i and j are connected. 2) Note that constraint for each alter is a part of a final matrix obtained from matrix calculation reflecting all edges. Ego has a maximum constraint of 0.44 but a minimum efficiency of 3/6 for alter E since E has the most indirect paths from ego to other alters.

order citations among cited inventions (alters-to-alters citations). The *ego-to-alters* citations simply identify a list of previous patents considered by an inventor, whereas the *alters-to-alters* citations characterize the network feature of cited inventions, considering not only the alters attributes but also their relational features. The 2nd order citations (alters-to-alters citations) play a critical role in characterizing the backward citation network structure in our study.[4]

Fig. 1 presents three examples of ego networks. Notice that they are citing the same number of previous patents, but the alters in each network are connected in different ways: alters in the left-most network (A) are highly clustered, whereas the alters in the right-most network (C) are completely isolated. To capture the differences in ego-network structure, we formally introduce three network indices: *constraint, cohesion,* and *efficiency*.

3.3.1. *Constraint*

The constraint is a dyadic index defined for each alter linked with an ego. This indicator measures the extent to which an ego's effort to maintain a relationship is concentrated on each alter. For an ego network, a network constraint is defined as an aggregation of the ego's dyadic constraints to each alter. Formally, an ego i's (patent i's) constraint can be defined as follows:

$$C_i = \sum_j \left(P_{ij} + \sum_q P_{iq} P_{qj} \right)^2 \tag{3}$$

Constraint consists of two components: direct and indirect networks. The direct constraint is represented solely by $\boldsymbol{P_{ij}}$, which denotes the proportion of i's relationship with j. In unweighted networks, it is assumed that i exerts the same effect on each of n alters, so $\boldsymbol{P_{ij}}$ becomes $1/n$. The indirect constraint represents the summation of all indirect paths that i can reach j through other alters. Therefore, an ego i's constraint on j becomes larger when j has more connections with other alters of i, which is denoted as subscript q in Eq. (3). For example, Fig. 2 demonstrates an ego network with six alters. Note that the constraint for each alter is a part of a matrix obtained from matrix calculation using all edges. In the network, the ego has a maximum constraint of 0.44 for alter E but a minimum constraint of 0.09 for alter A, since the ego's indirect links with the alters are maximized for E but minimized for A.

Constraint represents how much an ego's learning opportunities are restricted due to the network closure among alters. In Fig. 2, alter E has close relationships with other alters, which creates strain in learning from E due to other alters communicating with E. The amount of such strain would increase in proportion to the extent to which alters are clustered. In the network of backward citations, constraint reflects the degree of restriction imposed on prior knowledge in each citation as a candidate for recombination. Under the structural hole perspective, constrained networks are believed to have negative effects on performance. Relatedly, in a highly constrained ego-network of backward citations, it can be claimed that each knowledge source has a similar or redundant contribution to inventions, and thus, inventors have less

[4] To build and analyze ego networks, we use iGraph™, a graph analysis package used in the statistical software R.

opportunity to create novel ideas from them.

3.3.2. Cohesion

Cohesion measures the intensiveness of relationships among network members. For cohesiveness, we consider the clustering coefficient, which is widely used in conventional network analyses. The clustering coefficient is the ratio of actual ties that exist among alters to the maximum possible number of ties. The clustering coefficient can be calculated both at the network (global) and node (local) levels. The global clustering coefficient is the number of closed triplets (complete triangles) compared to the total number of triplets in the network. The local clustering coefficient is defined for each node and is given by the proportion of links between nodes within the focal node's neighbors divided by the maximum links that could exist among them (Watts and Strogatz, 1998). A node's local clustering coefficient, thus, becomes equivalent to the density of a network without the focal node. In an ego network, the local clustering coefficient is more relevant because it measures cohesion among alters from the perspective of ego. It is calculated as follows:

$$CC_i = \frac{t}{n(n-1)/2} \tag{4}$$

In Eq. (4), n is the number of alters and t the number of citations among them. The denominator is the maximum possible number of ties among n alters. Unlike the previous two measures aggregating measurements at the dyadic level, the clustering coefficient is defined for a focal node (i.e., ego). It measures the overall cohesiveness of an ego network. In the previous example in Fig. 2, the maximum possible ties among alters is 15 (= 6 × 5/2) and the actual number of ties among alters is 4. Hence, the clustering coefficient is 0.26 (= 4 /15).

In citations analyses, a high cohesion implies that there is pressure toward forming network closure among cited inventions. From the network closure perspective, a high cohesion leads to the greater homogeneity and uniformity among knowledge sources. This implies that knowledge in network closure is less distinctive but closely related or redundant with each other, and one could expect a greater dependence and homophiles among them. Therefore, one may claim that an invention created in a highly cohesive network of previous inventions is more likely to be incremental, but the invention likely has a lower technological risk and a higher possibility of working together with other technologies.

3.3.3. Efficiency

Efficiency represents the proportion of nonoverlapping knowledge that flows from alters to ego. If an alter can be reached via many other existing alters, a tie with the alter can be considered redundant, and thus, inefficient. Therefore, the actual amount of prior knowledge (i.e., the effective size of the network) that an ego effectively acquires must be reduced by the amount of redundancy. The effective size of a node i's ego network is measured as follows:

$$E_i = \sum_j \left[1 - \sum_q p_{iq}\, m_{qj}\right] \tag{5}$$

In Eq. (5), m_{qj} measures the amount of interaction between q and j divided by j's strongest interaction and represents the marginal strength of node q's relation with node j. For a binary network, like citation networks, every link has the same strength of 1, and thus, the value of m is 0 or 1 depending on whether or not j is connected to q. The summation, $\sum_q p_{iq}\, m_{qj}$, represents the redundancy of i's relation with j because it gets larger as i can reach j via other nodes. For each alter j, the effective size is 1 if the summation is 0.

Fig. 2 describes how effective size is calculated. For instance, the effective size of the tie with alter A remains the same as its original size 1 because A has no ties with other alters. However, the effective size of the tie with E is reduced by 3/6 (1–3/6), since there are 3 additional indirect links between E and ego. The sum of each tie's effective size becomes the effective size of the network, which is the number of direct ties to alters reduced by the amount of redundancy. Efficiency is the ratio of the effective size of the ego's network to its initial size.

The efficiency of the network normalizes the effective size of the ego network and tells us how much unique knowledge an ego obtains from each alters. If there are no ties among alters, the knowledge received from alters is completely not overlapped. In this case, the effective size of each alter is 1, and the effective size of the network is the number of alters. The associated efficiency of the network is equal to 1. In contrast, if all alters are completely tied to each other, the ties by the ego are completely ``redundant,'' and the ego obtains no unique knowledge from alters. In this case, the effective size of each alter is 1- $(n$-1$)/n$, and the effective size of the network is equal to 1. The associated efficiency of the network is equal to $1/n$. The definition of efficiency reveals the unique information existing in a network of cited inventions by detaching 'redundant' knowledge from the network.

3.3.4. Other variables for prior art characteristics

To supplement the network measures, we add several variables that capture additional information from the prior arts: *technology scope, temporal scope, and recency*. These variables are well-known factors in patent analyses, and their effects on patent values have been studied extensively in the literature. By controlling the variables, we can capture the net effects of network structure on patent value. *Technology scope* represents the extent to which backward citations are spread across various technology sectors. This is measured by the unique number of technology classes that each cited patent belongs to. *The temporal scope* is a measure for the yearly dispersion of backward citations and calculated by the standard deviation of citation lag years. *Recency* represents how recent a patent's referred knowledge is, and it is calculated as the mean lag years of backward citations (Lanjouw and Schankerman, 2004). Variables to control for the features unique to the patent itself are also considered. We include the number of claims (*# claims)*, the number of backward citations (*# backwards)*, and the number of technological classes that the patent belongs (*# classes)*. Year dummies and patent class dummies are also included in the regression models to control for the unobserved fixed effects.

3.4. applicant vs. examiner backward citations

As mentioned, recent studies have paid attention to the distinctive incentives of examiner citations compared with applicant citations (Alcácer et al., 2009). Examiners cite prior art that an inventor missed in reporting, and thus, tend to block or limit the claimed inventions. In this case, aggregate backward citations may not effectively represent knowledge spillover from prior art to an inventor. Considering this issue, applicant backward citations are employed to calculate network variables and other independent variables. By excluding examiner backward citations in measuring the prior art structure, the proposed variables can represent actual knowledge flows from the prior art.[5]

[5] It should be noted that the exclusion of examiner citations is only applied to the 1st order backward citations. Although the samples are applied after 2001, most of the 2nd order backward citations are applied before 2001 and examiner citations cannot be identified. Although this is due to data limitation, we believe that this would not affect the empirical results. The 1st order citations are directly cited and suitable for representing a knowledge flow to inventors. However, the 2nd order citations are not cited by applicants. They are more associated with technological relationships among cited inventions rather than knowledge flow to inventors.

Table 1
Summary Statistics.

Variables		Mean (S.E.)	Median	Min.	Max.	S.D.
Forward citations count	# of total forward citations	20.155 (0.332)	8.00	0.00	1449	40.54
	# of examiner citations	1.870 (0.026)	1.00	0.00	94.00	3.21
	Survival time of total forward citations	11.092 (0.025)	13.00	1.00	13.00	2.87
	Survival time of examiner citations	8.261 (0.037)	9.00	1.00	13.00	3.45
Network structure of backward citations	Constraint	0.076 (0.001)	0.07	0.01	0.43	0.04
	Cohesion	0.121 (0.001)	0.09	0.00	0.80	0.11
	Efficiency	0.892 (0.001)	0.93	0.12	1.00	0.11
Diversity and recency of backward citations	# of backward citations	23.194 (0.130)	17.00	10.00	97.00	15.91
	Technology scope: *Unique # of backward cited classes*	4.326 (0.011)	4.00	1.00	10.00	1.35
	Temporal scope: *Standard deviation of citation lag year*	6.763 (0.036)	5.84	0.18	58.65	4.35
	Recency: *Mean of backward citation lag years*	11.228 (0.044)	10.14	1.58	21.25	5.34
Patent information	# of clams	23.097 (0.186)	18.00	1.00	513	22.70
	# of classes	1.675 (0.006)	2.00	1.00	6.00	0.77
	Years to expiration (expired patents)	8.67 (0.024)	8.01	3.5	12	
	Renewal probability (1 – the proportion of expired patents)	43.95%				

However, it is also important to notice that applicant citation is far from being a perfect tool for representing the exact structure of relevant prior inventions (Jaffe et al., 2000). According to Alcácer et al. (2009), the share of examiner citations is significant, as 41% of all citations on U. S. patents in 2001–2003 came from examiners rather than applicants. In this case, the comprehensive structure of previous inventions could be identified better with aggregate backward citations. This is because examiners, as experts in the field, add a significant portion of backward citations considered as relevant prior art to the given invention. The examiner-added backward citations limit the scope of the claimed inventions and play a key role in determining the patentability of an invention. As the main concern of the study is to examine whether the network structure of backward citations provides any meaningful information on the future patent value, we also employ aggregate backward citations, including examiner citations, and compare the results with that using applicant backward citations only.

4. Empirical results

4.1. Data and summary statistics

We select the U.S. pharmaceutical patents granted in 2001∼2005 for our analysis. The pharmaceutical sector has been recognized as an industry that relies heavily on legal protection systems for appropriation and has been one of the most studied sectors in the related literature (Cohen et al., 2000). Among the pharmaceutical patents identified by the NBER patent classification scheme (Hall et al., 2001), all patents granted during 2001∼2005 are considered for the research sample.[6] Although the samples are from 2001 to 2005, building citation networks requires the entire universe of citations. We utilize a comprehensive patent dataset from 1975 to 2018 to construct the ego-network of backward citations for each sample patent.

From the initial candidates of 47,076 pharmaceutical patents, a substantial number of patents were excluded due to missing information. The patents citing old patents granted prior to 1975 were also excluded because the 2nd order citations are not available. The patents having few backward citations (less than 5) were also excluded because network variables for such patents have little variability, and hence, are not directly comparable with other patents.[7] Consequently, the final sample includes 13,917 U.S. pharmaceutical patents.

Table 1 presents the descriptive statistics for the variables used in the analysis. On average, the sample patents have 23.097 claims, 23.194 total backward citations, and belong to less than 2 technological classes. The sample patents also cite 4.326 technology classes (fields) on average, which indicates that roughly 5.36 (23.194 / 4.326) patents are cited in each technological field. Regarding the temporal aspect of citations, the mean of backward citation lags (*recency*) is 11.228 years, meaning that, on average, about 11-year-old patents are cited in the pharmaceutical R&D. The sample patents are cited 20.155 (*forward citations*) times on average, while many patents received no citations at all. Forward citations have a right-skewed distribution in which the mean (20.155) is much larger than the median (8.00).

As mentioned in the previous section, the number of examiner forward citations is considered for a dependent variable to represent the private value of a patent. The average number of examiner forward citations was much smaller than that of the total forward citations (1.870 vs. 20.155), which indicates that most forward citations originate from applicants. The mean survival time for total forward citations is 11.092

[6] The subcategory 'drugs' in the NBER patent classifications is used to define the pharmaceutical sector. This category corresponds to the USPTO class codes 424 and 514 (Drug, bio-affecting and body treating compositions).

[7] We conducted the same set of regression analyses using the patent data set, which has more than or equal to 2 backward citations. The regression results were quite consistent with the original version of regression results, which utilizes patent data set with more than or equal to 5 backward citations.

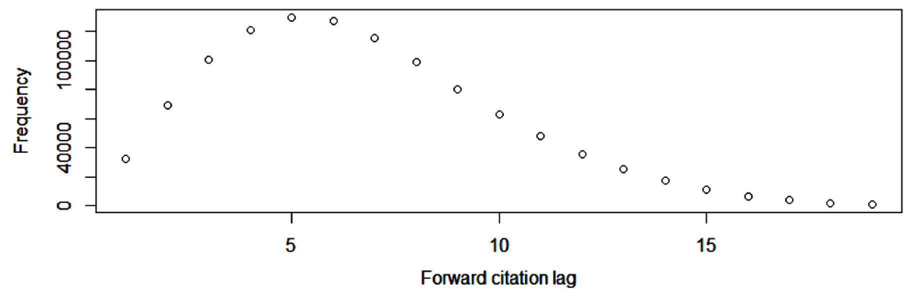


Fig. 3. Distribution of Forward Citations in the U.S. Pharmaceutical Industry.

Table 2
Examiner Forward Citations: Negative Binomial Regression.

	Dependent Variable: Examiner forward citations							
	Applicant backward citations				*Aggregate backward citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constraint		−7.469***				−5.046***		
		(0.837)				(0.667)		
cohesion			−1.198***				−0.844***	
			(0.150)				(0.121)	
efficiency				1.362***				1.511***
				(0.163)				(0.134)
tech. scope	0.590***	0.528***	0.545***	0.535***	0.134***	0.125***	0.127***	0.117***
	(0.055)	(0.054)	(0.054)	(0.054)	(0.011)	(0.011)	(0.011)	(0.011)
temp. scope	0.033***	0.034***	0.034***	0.033***	0.028***	0.030***	0.031***	0.029***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
recency	−0.043***	−0.043***	−0.045***	−0.044***	−0.037***	−0.039***	−0.037***	−0.038***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
# claims	0.294***	0.285***	0.291***	0.289***	0.280***	0.279***	0.278***	0.274***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)	(0.015)
# classes	0.302***	0.308***	0.301***	0.302***	0.173***	0.175***	0.177***	0.176***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.016)	(0.016)	(0.016)	(0.016)
# citations	0.070***	−0.223***	0.057**	0.059**	0.029	0.019	−0.179***	0.021
	(0.025)	(0.042)	(0.025)	(0.025)	(0.023)	(0.023)	(0.036)	(0.023)
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
patent class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,289	10,289	10,289	10,289	13,697	13,697	13,697	13,697
Log Likelihood	−19,725	−19,636	−19,661	−19,654	−25,357	−25,309	−25,297	−25,233
Akaike Inf. Crit.	39,473	39,297	39,346	39,333	50,740	50,647	50,622	50,495

* $p<0.1$; ** $p<0.05$; *** $p<0.01$.

and 8.261 for examiner forward citations. This implies that examiners tend to value and cite more recent technologies than applicants, which shows consistent results with Alcácer and Gittelman (2006). Finally, the mean of renewal dummy (43.95%) indicates that 43.95% of patents had not expired up to the third renewal due date of maintenance (i.e., the 12th year after registration).

It is noteworthy that the analysis imposes a fixed time window to address the truncation problem associated with measuring forward citations and renewal probability. Truncation is a typical issue that occurs in citation analyses because recent patents, with all other things being equal, have less chance of being cited compared to older patents. The same problem happens with renewal probability since recent patents have a lesser probability of expiration. Fig. 3, the histogram of forward citation lags, shows that the forward citations almost fade out within about 15 years after being granted. Given that the last year of the sample is 2005 and the last year of citation records is 2018, we impose a 13-year fixed time window in counting forward citations for all the sample patents.

4.2. *Effects on private value*

Tables 2, 3, and 4 present the regression results on how the network structure of backward citations affect the private value of an invention. Following the previous studies, we measure the private value of invention using 'the count of examiner forward citations' and 'a dummy for patent renewal.' Tables 2 and 3 display the empirical results with 'examiner forward citations' as dependent variables, and Table 4 for patent renewal. The primary results are obtained employing only applicant backward citations for measurement to characterize the knowledge flow to a given invention, and the comparable set of results was obtained from the same analyses using aggregated backward citations. The left four columns (1)-(4) show the empirical results using only applicant backward citations, and the right four columns (5)-(8) display the results in which aggregate backward citations are used. Year dummies and patent class dummies are used in all regression analyses.

Table 2 presents the results of the negative binomial estimation. Models (1) and (5) show the estimation results without network variables, while the remaining models include various network variables. In all model specifications, the controls show consistent results with previous studies. The number of claims (*# claims*) and the number of classes (*# classes*) are positively correlated with examiner forward citations, suggesting that the breadth of patent claims, controlling for other factors, positively affects the private value of an invention. The number of backward citations (*# citations*) also shows a positive correlation with

Table 3
Examiner Forward Citations: Cox Hazard Ratio.

	Dependent variable: Cox hazard of examiner forward citation							
	Applicant backward citations				*Aggregate backward citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constraint		4.732***				3.460***		
		(0.574)				(0.566)		
cohesion			0.980***				1.092***	
			(0.107)				(0.107)	
efficiency				−0.975***				−0.875***
				(0.114)				(0.119)
tech. scope	−0.340***	−0.303***	−0.305***	−0.302***	−0.071***	−0.060***	−0.066***	−0.062***
	(0.039)	(0.039)	(0.039)	(0.039)	(0.010)	(0.010)	(0.010)	(0.010)
temp. scope	−0.015***	−0.017***	−0.017***	−0.016***	−0.006	−0.008*	−0.007*	−0.007
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
recency	0.015***	0.016***	0.017***	0.016***	0.001	0.004	0.001	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
# claims	−0.197***	−0.195***	−0.196***	−0.194***	−0.085***	−0.085***	−0.083***	−0.083***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
# classes	−0.129***	−0.134***	−0.130***	−0.130***	−0.038***	−0.041***	−0.042***	−0.040***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.014)	(0.014)	(0.014)	(0.014)
# citations	−0.033	0.194***	−0.021	−0.024	0.037*	0.053**	0.180***	0.045**
	(0.023)	(0.036)	(0.023)	(0.023)	(0.021)	(0.021)	(0.031)	(0.021)
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
patent class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,289	10,289	10,289	10,289	8565	8565	8565	8565
R^2	0.071	0.077	0.079	0.078	0.033	0.044	0.038	0.038
Wald Test	762.3***	832.9***	848.8***	836.4***	307.780***	409.510***	350.130***	349.600***
	(df = 12)	(df = 13)	(df = 13)	(df = 13)	(df = 12)	(df = 13)	(df = 13)	(df = 13)

*$p<0.1$; **$p<0.05$; ***$p<0.01$.

Table 4
Patent Renewal: Logistic Regression.

	Dependent variable: Renewal probability							
	Applicant backward citations				*Aggregate backward citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constraint		−2.609**				−2.268**		
		(1.136)						
						(0.951)		
cohesion			−0.664***				−0.533***	
			(0.200)				(0.173)	
efficiency				0.523**				0.546***
				(0.214)				(0.186)
tech. scope	0.494***	0.470***	0.465***	0.470***	0.408***	0.385***	0.384***	0.381***
	(0.073)	(0.074)	(0.074)	(0.074)	(0.067)	(0.068)	(0.068)	(0.068)
temp. scope	0.020***	0.021***	0.021***	0.020***	0.020***	0.021***	0.021***	0.020***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
recency	−0.039***	−0.039***	−0.040***	−0.040***	−0.036***	−0.036***	−0.037***	−0.037***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
# claims	0.241***	0.239***	0.239***	0.239***	0.229***	0.228***	0.229***	0.227***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.022)	(0.022)	(0.022)	(0.022)
# classes	0.050	0.054	0.052	0.052	0.031	0.036	0.036	0.036
	(0.048)	(0.048)	(0.048)	(0.048)	(0.044)	(0.044)	(0.044)	(0.044)
# citations	−0.028	−0.151**	−0.032	−0.030	−0.016	−0.129**	−0.021	−0.019
	(0.040)	(0.067)	(0.040)	(0.040)	(0.038)	(0.060)	(0.038)	(0.038)
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
patent class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,289	10,289	10,289	10,289	13,697	13,697	13,697	13,697
Akaike Inf. Crit./ chi^2	13,500	13,496	13,491	13,496	15,952	15,948	15,944	15,945

*$p<0.1$; **$p<0.05$; ***$p<0.01$.

examiner forward citations, but it is significant only in Model (1) in which applicant backward citations are used. In addition, *technology scope* (the number of backward citation classes) and *temporal scope* (the dispersion of the backward citation lags) are also positively correlated with examiner forward citations. This implies that a patent's private value is proportional to both the dispersion of the citation lags and the breadth of the technological scope being cited. These variables are related to the extent of the technological search scope, and the results suggest that patents citing the broader scope in both technological and temporal aspects can create greater value than patents focused on a narrower scope. *Recency* is the mean of citation lags (so the smaller, the more recent) and shows a negative correlation with examiner forward citations. This finding suggests that, on average, recent knowledge is more preferred than older knowledge to create inventions of greater

Table 5
Total Forward Citation: Negative Binomial Regression.

	Dependent variable: Total forward citations							
	Applicant backward citations				*Aggregated backward citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constraint		−8.222***				−7.745***		
		(0.874)				(0.801)		
cohesion			0.202				0.103	
			(0.157)				(0.147)	
efficiency				0.061				0.339**
				(0.169)				(0.159)
tech. scope	0.394***	0.326***	0.403***	0.392***	0.091***	0.093***	0.082***	0.087***
	(0.058)	(0.059)	(0.058)	(0.059)	(0.014)	(0.014)	(0.014)	(0.014)
temp. scope	0.021***	0.021***	0.020***	0.021***	0.015***	0.015**	0.017***	0.016***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
recency	−0.040***	−0.038***	−0.039***	−0.040***	−0.039***	−0.038***	−0.039***	−0.040***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
# claims	0.286***	0.274***	0.287***	0.286***	0.267***	0.268***	0.262***	0.264***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)
# classes	0.340***	0.351***	0.341***	0.340***	0.179***	0.179***	0.184***	0.180***
	(0.037)	(0.038)	(0.037)	(0.037)	(0.021)	(0.021)	(0.021)	(0.021)
# citations	0.290***	−0.026	0.291***	0.289***	0.296***	0.296***	−0.022	0.296***
	(0.028)	(0.046)	(0.028)	(0.028)	(0.030)	(0.030)	(0.045)	(0.030)
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
patent class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,289	10,289	10,289	10,289	13,697	13,697	13,697	13,697
Log Likelihood	−45,728	−45,502	−45,724	−45,728	−59,030	−59,029	−58,741	−59,016
Akaike Inf. Crit.	91,479	91,028	91,473	91,481	118,087	118,086	117,512	118,061

*$p<0.1$; **$p<0.05$; ***$p<0.01$.

Table 6
Total Forward Citation: Cox Hazard Ratio.

	Dependent variable: Cox hazard of total forward citation							
	Applicant backward citations				*Aggregated backward citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constraint		6.681***				2.060***		
		(0.860)				(0.405)		
cohesion			0.446**				0.131*	
			(0.173)				(0.079)	
efficiency				−0.504***				−0.098
				(0.185)				(0.085)
tech. scope	−0.496***	−0.438***	−0.479***	−0.474***	−0.057***	−0.056***	−0.054***	−0.056***
	(0.061)	(0.062)	(0.062)	(0.062)	(0.008)	(0.008)	(0.008)	(0.008)
temp. scope	−0.012**	−0.015**	−0.013**	−0.013**	−0.004	−0.004	−0.005	−0.004
	(0.006)	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)	(0.003)
recency	0.016***	0.017***	0.017***	0.017***	0.011***	0.012***	0.011***	0.012***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
# claims	−0.279***	−0.272***	−0.278***	−0.277***	−0.076***	−0.075***	−0.075***	−0.075***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.010)	(0.010)	(0.010)	(0.010)
# classes	−0.205***	−0.217***	−0.208***	−0.208***	−0.046***	−0.047***	−0.049***	−0.047***
	(0.045)	(0.045)	(0.045)	(0.045)	(0.012)	(0.012)	(0.012)	(0.012)
# citations	−0.255***	0.084	−0.252***	−0.253***	−0.042**	−0.041**	0.067**	−0.042**
	(0.039)	(0.059)	(0.039)	(0.039)	(0.018)	(0.018)	(0.028)	(0.018)
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
patent class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,289	10,289	10,289	10,289	13,062	13,062	13,062	13,062
R^2	0.065	0.070	0.065	0.065	0.038	0.038	0.040	0.038
Wald Test	698.6***	771.5.3***	705.8***	706.1***	523.520***	526.730***	552.690***	525.040***
	(df = 12)	(df = 13)	(df = 13)	(df = 13)	(df = 12)	(df = 13)	(df = 13)	(df = 13)

*$p<0.1$; **$p<0.05$; ***$p<0.01$.

value.

Our primary interest is in the effects of network variables in Models (2)–(4) and (6)–(8). Overall, the effects of the network structure of the prior art on examiner forward citations are quite consistent and significant at 1% significance level. The two network variables, *constraint* and *cohesion,* show a statistically significant and negative correlation with examiner forward citations, while the *efficiency* of a network is positively correlated. For example, with all other factors being equal, a 1% increase in *constraint* tends to decrease the expected count of examiner forward citations by 7.20%. This implies that the network characteristics of the prior art structure (measured by applicant backward citations) has a significant correlation with the *ex-post* private value of an invention. That is, patents based on the more constrained, the more cohesive, or the less efficient structure of prior art are likely to have lower future private value. The analyses results employing aggregate backward citations in Models (6)–(8), in which we assume that aggregate backward citations provide more comprehensive information in identifying the strategic position of a patent, also show qualitatively similar results with

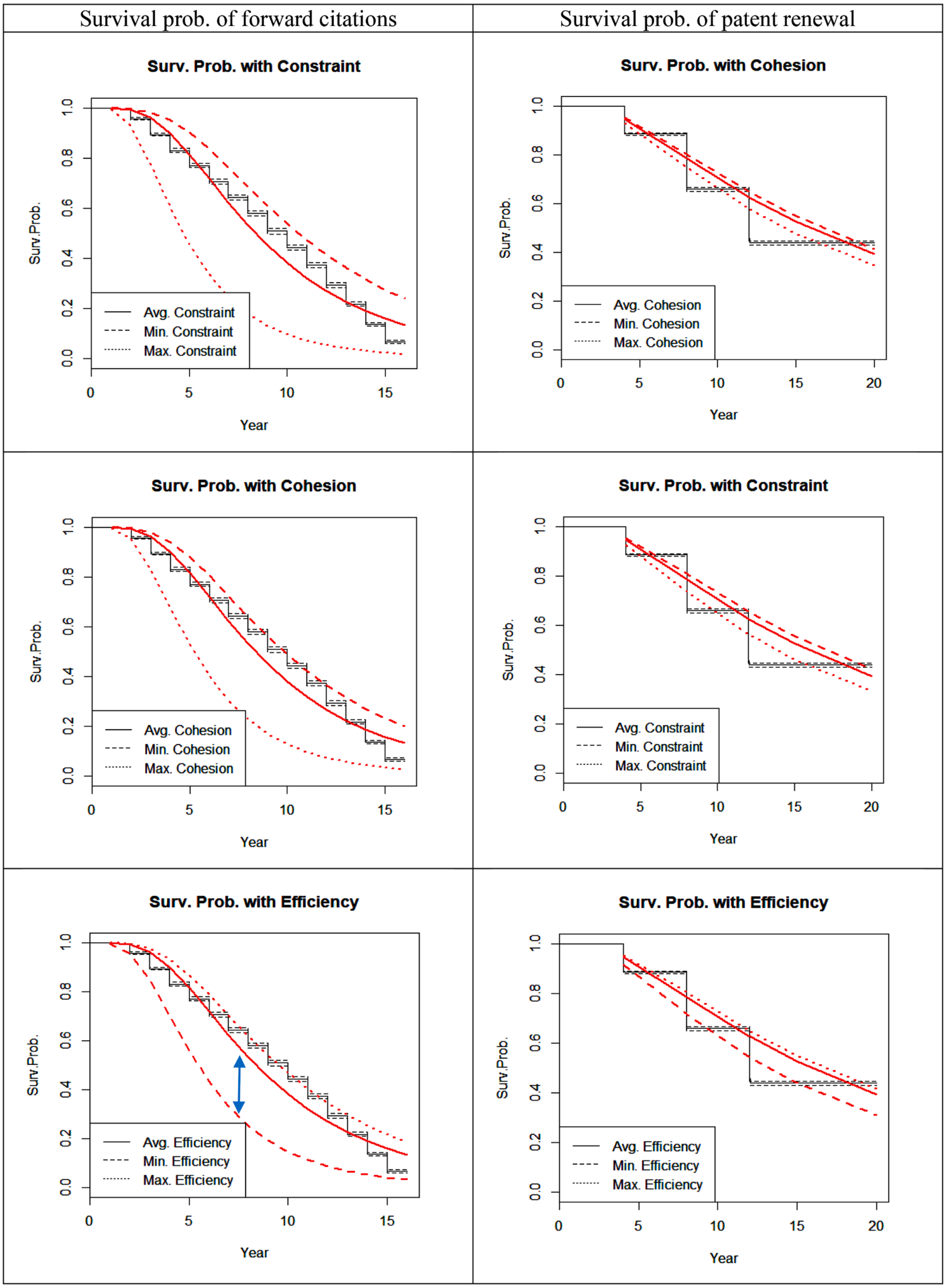


Fig. 4. Survival Probability with Network Variables. Notes: The stepped-lines represent a Kaplan-Meier curve. The solid, dashed, and dotted lines indicate the predicted survival probability for patents whose covariates are set to their average, maximum, and minimum values, respectively.

those ones using only applicant backward citations.

Table 3 presents the empirical results using Cox hazard regression analysis. In all model specifications, all network variables show statistically significant effects on the examiner forward citations. The two network variables, *constraint* and *cohesion,* show statistical significance at 1% level and positive correlation with the hazard ratio of examiner forward citations, while the *efficiency* of network shows negatively correlated estimates. For instance, *constraint* increases the deprecation rate of examiner forward citations by 4.85%. This implies that the hazard of citation termination at the next period increases with patents based on constrained and cohesive networks of previous inventions, while it decreases with efficient networks. The regression results from Tables 2 and 3 suggest that the hidden feature of the backward citation network provides important information on the future private value of a patent, not only for the level of patent value but also for its durability.

Finally, we test the effects of network variables on the patent *renewal probability*, which is another proxy for the private value of the invention. The renewal dummy variable represents whether a patent is in a state of expiry after the third due date of maintenance fee plus the 6-month grace period, which is the twelfth year after the date of issue.[8] Table 4 presents the logistic regression analyses for patent renewal. The overall results for patent renewal analyses are quite consistent with the previous results using examiner forward citations as dependent variables. For instance, the results show that a 1% increase in *constraint* decreases the odds of renewal by 2.58%. This implies that other factors being controlled, inventions based on the cohesive knowledge structure of prior art are likely to have less private value than others.

4.3. Effects on technological spillover

Tables 5 and 6 present the regression results using total forward examination as a dependent variable, which measures social impact or technological spillover of the invention. The overall effects of the backward citation network structure on the total forward citations are also consistent, but somewhat modest, depending on the network variables, in the negative binomial regression. In Models (2) and (6), *Constraint* shows a statistically significant negative effect on total forward citations, while *Cohesion* does not have statistical significance in Models (3) and (7). *Efficiency* shows a positively significant effect on total forward citations only in Model (8).

Table 6 presents the regression results of a Cox-hazard estimation and shows much stronger and clearer empirical results. In Models (2)–(4) where applicant backward citations are employed, all network variables show statistically significant effects on the hazard ratio of citation termination for total forward citations. Both *constraint* and *cohesion* show significant positive effects, while *efficiency* has a negative effect. This implies that the less constrained (and the less cohesive) and more efficient is the network structure of prior knowledge, the more impactful is the future technology spillover or value of the invention. Once again, the results suggest that the network structure of backward citations has significant implications on the social value of a patent, and the future impact of patents is likely to be less and depreciate faster when they are based on highly *constrained* and/or *cohesive* prior knowledge.

Although a few estimates of network variables lost statistical significance, it is remarkable that the signs of the coefficients are quite consistent with the previous models. The empirical results so far clearly suggest that the network structure of backward citations has important implications for the predictability of future patent value in both level and depreciation rates.

We also find that the empirical results support the structural hole perspective, which acknowledges the opportunity of brokerage in the gap between disparate knowledge sources. According to the structural hole perspective, a patent positioned at a structural hole in citation networks weaves knowledge from different fields and is more likely to create value from the novel recombination of distinct technology assessing a broader array of ideas.

It is important to note, however, that the interpretation of the empirical results should consider that we cannot necessarily claim that the structural hole strategy is the better or healthier patent development strategy than the network closure strategy. An innovation strategy is, in general, chosen considering a variety of factors with various circumstances, and patent application is a consequence of such strategic choices in which an inventor wants to realize the maximum return from the project. In this case, the optimal choice of innovation project could be either a ‘major and drastic’ invention or a ‘minor and incremental’ one. Invention under the structural hole strategy may have a higher private value or higher technological impact, but it also could have a higher risk of failure. To provide a more comprehensive answer to this question, we may need more information on the inventor’s effort, including their failed projects as well. The empirical findings, however, at least evidence that the network structure of backward citation has a significant correlation with the value of a patent and reveal important information on the predictability of patent value. In addition, the structural hole strategy creates a greater value of an invention in both private and social terms as long as a given invention is successful enough to register its patent.

Fig. 4 summarizes the main findings from the regression analyses. The left panels of the figure present the survival plot with forward citations, and the right panels present that with renewal probability. In the figure, the stepped-line in each graph represents a Kaplan-Meier curve, which indicates a non-parametric estimation of survival probability without covariates for the whole population. The other lines represent the estimates from survival models with a log-normal probability distribution. The solid line predicts the survival probability for the average patent in which each covariate is set to its mean value. The dashed line in each graph indicates the predicted citation survival probabilities and renewal probabilities for patents having a minimum value of each network variable, and the dotted lines represent the case having the maximum value for network variables. The figures show that the dotted lines for cohesion and constraint are located below the average curve, while the dotted line for efficiency is located above the average line. These effects are substantial. For example, in the third left panel of Fig. 4, the survival probability of forward citations at year 7 for the patent having the least efficient backward citation network is approximately 30%, compared to 60% for the patents with average efficiency. These results imply that the patents based on highly cohesive, highly constrained, and less efficient backward citation structures are likely to have the least technological impact and economic value.

Finally, we supplement the proportional hazard estimation with a survival probability estimation under a log-normal distribution and is presented in the Appendix. Notice that the signs of the network variables are opposite to those of Tables 3 and 6 because a Cox-hazard estimation uses a hazard ratio, while the log-normal survival analysis uses survival probability as a dependent variable. Similar and consistent results are obtained from the survival analysis using a log-normal estimation. All network variables are statistically significant at 1% significance level.

5. Conclusions and discussions

Modern innovation is a process of ‘search and recombination,’ and the value of an invention is associated with not only how many prior inventions are considered, but also how they are related to each other.

We employ social network analysis, broadly used in social capital theory, to investigate whether the hidden network structure of prior art has significant implications for future patent value. We consider

[8] Patent owners are required to pay periodic maintenance fees to preserve their patent rights. After a patent is issued, maintenance fees are payable to the USPTO at 3 specific time points: at 3 to 3.5 years, 7 to 7.5 years, and 11 to 11.5 years after the date of issue. If they fail to pay the fees within the due date, the patents will be expired.

network variables such as *constraint, cohesion, and efficiency* using both the 1st order (ego-to-alter) and the 2nd order citation (alter-to-alter) networks, and test whether these network characteristics reveal any statistically significant correlation with the future value of an invention. The backward citation network reveals key attributes regarding the knowledge structure of the prior art. In patent citation networks, 'constraint' reflects the degree of restriction imposed on prior knowledge as a candidate for recombination. 'Cohesion' measures the degree of relationships among network members, and implies that there is pressure toward forming greater homogeneity among knowledge sources. 'Efficiency' measures the proportion of nonoverlapping knowledge flows from alters to ego and determines how much unique knowledge an invention obtains from each unit in the prior arts. Following the previous studies, we assume that the network variables measured by applicant backward citation characterize the knowledge flow of an invention. Recognizing the bounded capability of inventors to list all relevant prior art and the complementary role of examiners adding the missed prior art, we also employ the variables measured by aggregate backward citations. We expect that the variables identify the strategic position of an invention with a more comprehensive information on the technological trajectory of prior art.

Using U.S. pharmaceutical patents data, we find that the network feature of backward citations has significant implications on the ex-post value of an invention in both levels and its longevity, and in both private value and social impact. Patents with more constrained, more cohesive, and less efficient backward citation networks are likely to have lower private patent value and lower technological impact. Given the condition that the invention is successful enough to register a patent, these results support the 'structural hole' hypothesis between the two contrasting perspectives on social capital: the network closure versus the structural hole perspectives. Under the structural hole hypothesis, the highly constrained, cohesive, and non-efficient citations may have similar or redundant contributions for an invention, having less opportunity to create novel ideas from them.

This study contributes to the literature of technology evaluation and extends the dimensions of analysis for patent evaluation by employing the second-order backward citation networks. Also, to the best of our knowledge, this study provides the first empirical evidence on the relationship between the various network measures of backward citations and the future value of a patent. We also suggest that our model is valuable for both academic and practical purposes. Companies and policymakers may apply our model to identify target technologies or to predict promising technological areas. With this model, companies also can choose target patents they want to purchase or R&D projects to develop against their competitor's patents. This strategic positioning of a company's R&D investment using network analysis will be the next step in our future research.

CRediT authorship contribution statement

Wonchang Hur: Software, Formal analysis, Data curation, Methodology. **Junbyoung Oh:** Conceptualization, Validation, Investigation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgements

We are grateful to the editor and two anonymous referees for useful comments and suggestions. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (Social Science of Korea, NRF-2016S1A3A2923769).

Appendix A1

Table A1

Table A1
Examiner Forward Citation: Log-normal survival.

	Dependent variable: Survival probability of examiner forward citation							
	Applicant backward citations				*Aggregate backward citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constraint		−6.402***				−2.955***		
		(0.758)				(0.319)		
cohesion			−1.114***				−0.850***	
			(0.134)				(0.060)	
efficiency				1.115***				0.655***
				(0.144)				(0.066)
tech. scope	0.461***	0.401***	0.413***	0.409***	0.044***	0.035***	0.040***	0.037***
	(0.049)	(0.050)	(0.050)	(0.050)	(0.005)	(0.005)	(0.005)	(0.006)
temp. scope	0.019***	0.022***	0.021***	0.020***	0.003	0.005**	0.004*	0.004
	(0.005)	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
recency	−0.022***	−0.023***	−0.024***	−0.023***	0.0001	−0.002	0.00001	−0.0004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
# claims	0.262***	0.257***	0.259***	0.258***	0.057***	0.057***	0.057***	0.055***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.007)	(0.007)	(0.007)	(0.007)
# classes	0.201***	0.211***	0.204***	0.206***	0.030***	0.032***	0.033***	0.032***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.008)	(0.008)	(0.008)	(0.008)
# citations	0.049*	−0.256***	0.041	0.045	−0.015	−0.026**	−0.139***	−0.019
	(0.028)	(0.045)	(0.028)	(0.028)	(0.012)	(0.012)	(0.018)	(0.012)
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
patent class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,289	10,289	10,289	10,289	8565	8565	8565	8565
Chi^2	784.4***	855.5***	853.2***	844.1***	20.190***	502.983***	397.838***	397.288***
	(df = 13)	(df = 13)	(df = 13)	(df = 13)	(df = 12)	(df = 13)	(df = 13)	(df = 13)

*$p<0.1$; **$p<0.05$; ***$p<0.01$.

Appendix A2

Table A2

Table A2
Total Forward Citation: Log-normal survival.

	Dependent variable: Survival probability of total forward citation							
	Applicant backward citations				*Aggregate backward citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constraint		−6.429***				−1.092***		
		(0.646)				(0.166)		
cohesion			−0.279**				−0.182***	
			(0.117)				(0.032)	
efficiency				0.371***				0.110***
				(0.125)				(0.034)
tech. scope	0.340***	0.279***	0.328***	0.322***	0.027***	0.025***	0.025***	0.026***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.003)	(0.003)	(0.003)	(0.003)
temp. scope	0.017***	0.020***	0.017***	0.017***	0.002	0.002*	0.002*	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
recency	−0.023***	−0.024***	−0.023***	−0.023***	−0.004***	−0.004***	−0.004***	−0.004***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
# claims	0.214***	0.210***	0.214***	0.213***	0.037***	0.036***	0.036***	0.036***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.004)	(0.004)	(0.004)	(0.004)
# classes	0.163***	0.173***	0.164***	0.165***	0.028***	0.029***	0.030***	0.029***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.005)	(0.005)	(0.005)	(0.005)
# citations	0.167***	−0.144***	0.165***	0.166***	0.006	0.005	−0.052***	0.006
	(0.024)	(0.039)	(0.024)	(0.024)	(0.007)	(0.007)	(0.011)	(0.007)
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
patent class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,289	10,289	10,289	10,289	13,062	13,062	13,062	13,062
Chi^2	1162.1***	1260.8***	1167.8***	1170.9***	706.743***	739.818***	750.167***	717.139***
	(df = 12)	(df = 13)	(df = 13)	(df = 13)	(df = 12)	(df = 13)	(df = 13)	(df = 13)

*p<0.1; **p<0.05; ***p<0.01.

References

Ahuja, G., 2000. Collaboration networks, structural holes, and innovation: a longitudinal study. Adm. Sci. Q. 45 (3), 425–455.
Ahuja, G., Lampert, C.M., 2001. Entrepreneurship in the large corporation: a longitudinal study of how established firms create breakthrough inventions. Strateg. Manage. J. 22 (6–7), 521–543.
Albert, M.B., Avery, D., Narin, F., McAllister, P., 1991. Direct validation of citation counts as indicators of industrially important patents. Res. Policy 20 (3), 251–259.
Alcácer, J., Gittelman, M., 2006. Patent citations and a measure of knowledge flow: the influence of examiner citations. Rev. Econ. Stat. 88 (4), 774–779.
Alcácer, J., Gittelman, M., Sampat, B., 2009. Applicant and examiner citations in U.S. patents: an overview and analysis. Res. Policy 38, 415–427.
Allison, J.R., Lemley, M.A., 2000. Who's patenting what? An empirical exploration of patent prosecution. Vanderbilt Law Rev. 53, 2099–2180.
Arthur, W.B., 2007. The structure of invention. Res. Policy 36, 274–287.
Battkea, B., Schmidt, T.S., Stollenwerka, S., Hoffmann, V.H., 2016. Internal or external spillovers-which kind of knowledge is more likely to flow within or across technologies. Res. Policy 45, 27–41.
Burt, R.S., 1992. Structural Holes. Harvard University Press, Cambridge, MA.
Burt, R.S., 2004. Structural holes and good ideas. Am. J. Sociol. 110 (2), 349–399.
Carpenter, M.P., Narin, F., Woolf, P., 1981. Citation rates to technologically important patents. World Patent Inf. 3 (4), 160–163.
Christensen, J.F., 2002. Building innovate assets and dynamic coherence in multi-technology companies. In: Foss, N.J., Robertson, P.L. (Eds.), Resource, Technology and Strategy. Explorations in the Resource-Based Perspective, 123–. Routledge, London, p. 152.
Caballero, R.J., Jaffe, A.B., 1993. How high are the giants' shoulders: an empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. NBER Macroecon. Annu. 8, 15–74.
Cohen, W., Levinthal, D., 1990. Absorptive capacity: a new perspective on learning and innovation. Adm. Sci. Q. 35, 128–152.
Cohen, W.M., Nelson, R.R., Walsh, J., 2000. In: Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent. Cambridge, MA. NBER Working Paper No. 7552.
Coleman, J.S., 1988. Social capital in the creation of human capital. Am. J. Sociol. 94, 95–120.
Collins, C.J., Clark, K.D., 2003. Strategic human resource practices, top management team social networks, and firm performance: the role of human resource practices in creating organizational competitive advantage. Acad. Manage. J. 46 (6), 740–751.
Corredoira, R.A., Banerjee, P.M., 2015. Measuring patent's influence on technological evolution: a study of knowledge spanning and subsequent inventive activity. Res. Policy 44 (2), 508–521.
Cotropia, C.A., Lemley, M.A., Sampat, B., 2013. Do applicant patent citations matter? Res. Policy 42 (4), 844–854.
Danneels, E., 2002. The dynamics of product innovation and firm competences. Strateg. Manage. J. 23, 1095–1121.
Dahlin, K.B., Behrens, D.M., 2005. When is an invention really radical? Res. Policy 34 (5), 717–737.
Dosi, G., 1982. Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. Res. Policy 11, 147–162.
Freeman, L.C., 1979. Centrality in social networks conceptual clarification. Soc Netw. 1 (3), 215–239.
Granovetter, M.S., 1973. The strength of weak ties. Am. J. Sociol. 78 (6), 1360–1380.
Granstrand, O., 1998. Towards a theory of technology-based firm. Res. Policy 27, 465–489.
Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2001. In: The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. NBER Working Paper 8498.
Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2005. Market value and patent citations. RAND J. Econ. 36 (1), 16–38.
Hargadon, A., Sutton, R.I., 1997. Technology brokering and innovation in a product development firm. Adm. Sci. Q. 42 (4), 716–749.
Harhoff, D., Narin, F., Scherer, F.M., Vopel, K., 1999. Citation frequency and the value of patented inventions. Rev. Econ. Stat. 81 (3), 511–515.
Harhoff, D., Scherer, F.M., Vopel, K., 2003. Citations, family size, opposition and the value of patent rights. Res. Policy 32 (8), 1343–1363.
Hedge, D., Sampat, B., 2009. Examiner citations, applicant citations, and private value of patents. Econ. Lett. 105 (3), 287–289.
Helfat, C.E., 1994. Evolutionary trajectories in petroleum firm R&D. Manage. Sci. 40 (12), 1720–1747.
Henderson, R., Cockburn, I., 1994. Measuring competence? Exploring firm effects in pharmaceutical research. Strateg. Manage. J. 15, 63–84.
Iansiti, M., West, J., 1997. Technology integration: turning great research into great products. Harv. Bus. Rev. 75, 69–79.
Jaffe, A.B., 1986. Technological opportunity and spillovers of R&D: evidence from firms' patents, profits, and market value. Am. Econ. Rev. 76 (3), 984–1001.
Jaffe, A.B., 1989. Characterizing the `technological position' of firms with application to quantifying technological opportunity and research spillovers. Res. Policy 18, 87–97.
Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Q. J. Econ. 108 (3), 577–598.
Jaffe, A.B., Trajtenberg, M., Fogarty, M.S., 2000. Knowledge spillovers and patent citations: evidence from a survey of inventors. Am. Econ. Rev. 90 (2), 215–218.
Jung, H.J., Lee, J.J., 2016. The quest for originality: a new typology of knowledge search and breakthrough inventions. Acad. Manage. J. 59 (5), 1725–1753.
Kogut, B., Zander, U., 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. Organ. Sci. 3 (3), 383–397.

Kuhn, J., Younge, K., Marco, A., 2020. Patent citations reexamined. RAND J. Econ. 51 (1), 109–132.
Lanjouw, J.O., Schankerman, M., 2001. Characteristics of patent litigation: a window on competition. RAND J. Econ. 32 (1), 129–151.
Lanjouw, J.O., Schankerman, M., 2004. Patent quality and research productivity: measuring innovation with multiple indicators. Econ. J. 114 (495), 441–465.
Lemley, M., 2001. Rational ignorance at the patent office. Northwest. Univ. Law Rev. 95, 1495–1532.
Lampe, R., 2012. Strategic citation. Rev. Econ. Stat. 94 (1), 320–333.
Lerner, J., 1994. The importance of patent scope: an empirical analysis. Rand. J. Econ. 25 (2), 319–333.
Levinthal, D.A., March, J.G., 1993. The myopia of learning. Strateg. Manage. J. 14, 95–112.
Martin, X., Mitchell, W., 1998. The influence of local search and performance heuristics on new design introduction in a new product market. Res. Policy 26 (7–8), 753–771.
Mehta, A., Rysman, M., Simcoe, T., 2010. Identifying the age profile of patent citations: new estimates of knowledge diffusion. J. Appl. Econom. 25 (7), 1179–1204.
Moser, P., Ohmstedt, J., Rhode, P.W., 2017. Patent citations—an analysis of quality differences and citing practices in hybrid corn. Manage. Sci. 64 (4), 1926–1940.
Nahapiet, J., Ghoshal, S., 1998. Social capital, intellectual capital, and the organizational advantage. Acad. Manage. Rev. 23 (2), 242–263.
Nelson, R., 1959. The simple economics of basic scientific research. J. Polit. Economy 67, 297–306.
Nemet, G.F., Johnsona, E., 2012. Do important inventions benefit from knowledge originating in other technological domains? Res. Policy 41, 190–200.
Nesta, L., Saviotti, P., 2005. Coherence of the knowledge base and firm innovative performance: evidence from the U.S. pharmaceutical industry. J. Ind. Econ. 53, 123–142.
Peng, T., Wang, Z.Z., 2013. Network closure, brokerage, and structural influence of journals: a longitudinal study of journal citation network in Internet research (2000–2010). Scientometrics 97 (3), 675–693.
Quintana-García, C., Benavides-Velasco, C.A., 2008. Innovative competence, exploration and exploitation: the influence of technological diversification. Res. Policy 37 (3), 492–507.
Reagans, R., McEvily, B., 2003. Network structure and knowledge transfer: the effects of cohesion and range. Adm. Sci. Q. 48 (2), 240–267.
Reagans, R., Zuckerman, E.W., 2001. Networks, diversity, and productivity: the social capital of corporate R&D teams. Organ. Sci. 12 (4), 502–517.
Rosenkopf, L., Nerkar, A., 2001. Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. Strateg. Manage. J. 22 (4), 287–306.
Sampat, B.N., Mowery, D.C., Ziedonis, A., 2003. Changes in university patent quality after the Bayh-Dole act: a re-examination. Int. J. Ind. Organ. 21 (9), 1371–1390.
Schoenmakers, W., Duysters, G., 2010. The technological origins of radical inventions. Res. Policy 39 (8), 1051–1059.
Savino, T., Petruzzelli, A.M., Albino, V., 2017. Search and recombination process to innovate: a review of the empirical evidence and a research agenda. Int. J. Manage. Rev. 19 (1), 54–75.
Shane, S., 2001. Technological opportunities and new firm creation. Manage. Sci. 47 (2), 205–220.
Silverman, B.S., 1999. Technological resources and the direction of corporate diversification: toward an integration of the resource-based view and transaction cost economics. Manage. Sci. 45 (8), 1109–1124.
Sorenson, J.B., Stuart, T.E., 2000. Aging, obsolescence and organizational innovation. Adm. Sci. Q. 45, 81–112.
Stuart, T.E., Podolny, J.M., 1996. Local search and the evolution of technological capabilities. Strateg. Manage. J. 17 (7), 21–38.
Subramaniam, M., Youndt, M.A., 2005. The influence of intellectual capital on the types of innovative capabilities. Acad. Manage. J. 48, 450–463.
Suzuki, J., Kodama, F., 2004. Technological diversity of persistent innovators in Japan. Res. Policy 33 (3), 531–549.
Teece, D.J., Pisano, G., Shuen, A., 1997. Dynamic capabilities and strategic management. Strateg. Manage. J. 18 (7), 509–533.
Tong, X., Frame, J.D., 1994. Measuring national technological performance with patent claims data. Res. Policy 23 (2), 133–141.
Trajtenberg, M., Henderson, R., Jaffe, A., 1997. University versus corporate patents: a window on the basicness of invention. Econ. Innov. New Technol. 5 (1), 19–50.
Uzzi, B., 1997. Social structure and competition in interfirm networks: the paradox of embeddedness. Adm. Sci. Q. 42 (1), 35–67.
Vasudeva, G., Zaheer, A., Hernandez, E., 2013. The embeddedness of networks: institutions, structural holes, and innovativeness in the fuel cell industry. Organ. Sci. 24 (3), 645–663.
Wang, J., Chiang, C., Lin, S., 2010. Network structure of innovation: can brokerage or closure predict patent quality? Scientometrics 84 (3), 735–748. https://doi.org/10.2224/sbp.2010.38.9.1229 https://doi.org/.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of `small-world' networks. Nature 393 (4), 440–442.