Innovation and Endogenous Knowledge Network Dynamics*

Jie Cai

School of Economics

Shanghai University of Finance and Economics

Shanghai, 200433, China

april.cai@gmail.com

Can Tian

Department of Economics

University of North Carolina at Chapel Hill

Chapel Hill, NC 27599, USA

tiancan@email.unc.edu

February 5, 2023

Abstract

We study firm innovation and path-dependent knowledge-network dynamics in a general-equilibrium framework with long-run growth. The empirical examination approximates knowledge flows by citations. It suggests that patent-level citation formation depends on patent quality and previous citations and that citation dynamics affect patenting firms' performance. A theoretical model is built to examine forward-looking firms' innovation incentives with path-dependent and payoff-relevant citations, where previous citations guide how each firm cites existing patents as inputs to produce new ones. The stationary equilibrium with balanced growth requires the patent-level citation-network formation process to be consistent with firm decisions. Path-dependent citations lead to another knowledge externality as older citations have spillover effects on newer ones' formation. Meanwhile, a patent may contribute to the value of every patent it cites as input, so newer patents' values may spill over to older ones. These externalities imply new challenges and also new possibilities for innovation policies.

JEL Classification codes: D21 D85 L14 O31 O34 O41

Keywords: innovation, patent citations, knowledge network formation, endogenous growth, R&D policies

*We thank Enghin Atalay, Santiago Caicedo, Mishel Ghassibe, Basile Grassi, Doug Hanley, Ernest Liu, Song Ma (discussant), Kris Nimark, Ezra Oberfield, Ben Pugsley, Andrey Stoyanov, Daniel Xu for valuable comments at various stages of this project. We also thank seminar and conference participants at University of Pittsburgh, SED, NASM, CICM, I85 Macro Workshop.

1 Introduction

Innovation is sequential. Innovators stand on the shoulders of giants, combining and improving existing knowledge to create something new and valuable. How does one find said giants? Avant-garde innovators discover new ways to use knowledge, which benefits subsequent innovators by inspiring them to follow similar paths. It is also true in academic research, as we often follow established citation paths to trace the previous literature when writing new articles. Therefore, knowledge externality means not only the height of giants' shoulders but also where to find them. Both kinds of spillovers affect innovators that are simultaneously knowledge users and producers, which impacts long-run economic growth and innovation policies.

This article studies directed and path-dependent knowledge flows and their impacts on innovation outcomes in a general-equilibrium framework. We begin with an empirical examination to show that it is plausible to describe knowledge flows approximated by patent citations as path- and quality-dependent and that citations are payoff-relevant to innovating firms. Then, we design a theoretical framework to highlight firms' innovative incentives with the path-dependent innovation process modeled as patent-level citation-network formation governed by firm decisions. The model is stylized and focuses on a stationary equilibrium with balanced growth. It captures that path-dependent innovation means additional knowledge externality, and an older patent's knowledge content spills over to newer ones citing it. When firms benefit from citations received, a newer patent contributes to the continuation values of all the older ones cited by bringing them additional citations in the future.

The empirical examination in Section 2 exploits patent citations from the U.S. Patent and Trademark Office (USPTO) and yields a set of new findings on patent-citation dynamics and their impacts on patenting firms. Citations are our proxy of directed knowledge-flow paths at the patent level. The first stylized fact implies that patent quality and previous citations jointly drive the citation formation. Furthermore, patents exhibiting more independence of previous citations tend to have higher quality. Here, the aggregate influence of previous citations on new ones is reflected by highly *transitive* citations. That is if patent b cites patent a, and patent c cites b, then it is also likely that c cites a. Then, we focus on the *first* citation ever observed from one technology category to another, named as a cross-category *pathway*. By definition, pathways capture new ways to use existing knowledge at the category level. We call the *citing* patent of a pathway a *pathfinder* and the *cited* patent a *path-receiver*. Relative to an average patent, pathfinders show more path independence and tend to score higher on widely-accepted quality measures. Meanwhile, path-receivers also have above-average quality, confirming quality-driven citations. The rest of the findings concern patenting firms, each of which may get involved in pathway creations directly as the owner firm of a pathfinder or receiver patent and indirectly as a peer firm innovating in the same category as the owner. We document a spillover effect of firms' innovation output in the citing category, such that after pathways appear, peer firms to pathfinder firms tend to innovate more. The last finding highlights the benefit of receiving citations, which spills over to peer firms, such that after pathways appear, path-receiver firms and their peers tend to show improved market performance.

The empirical findings have the following interpretations to motivate the model. First, regarding patent citation dynamics, the findings suggest that citation formation is not purely random but quality-directed and path-dependent. High-quality new patents likely make citations independent of previous citations, and citations likely go to high-quality existing patents. Therefore, citation dynamics are endogenous when patent quality, entry of new patents, and citation formation result from firm decisions. Second, regarding firm decisions, citation formation suggests a nontrivial selection of inputs for knowledge production and the possibility of following pre-existing paths. Furthermore, citations are payoff-relevant, such that firms benefit from receiving citations.

The theoretical model aims to capture these observations and examines the interdependence between firm innovation incentives and patent citation dynamics in general equilibrium. A continuum of infinitesimal firms own patents that differ in quality levels. Patent quality determines the profitability of the associated product, but it depreciates over time. Each firm hires researchers to innovate and decides the arrival rate and the expected quality of new patents. A new patent uses (cites) existing patents as inputs and improves upon the quality of cited inputs at the cost of research labor. A greater step size of quality improvement requires to hire more research labor and leads to better selections of input quality. That is, citing high-quality patents elevates new-patent quality but is costly. However, previously established citation edges may help to guide the input search and reduce the cost, as described by a citation process discussed later. As such, previous paths of knowledge flows directly impact innovation incentives and the formation of new citations, which, in turn, influence subsequent innovation. Furthermore, an existing patent accumulates citations at a rate consistent with the citation dynamics aggregated across all newer patents. If firms benefit from every citation received, the return to innovation includes a stream of citation values in addition to the product-market profit stream, both of which depend on the choice of new-patent quality. In this case, citation dynamics become an equilibrium object, as forward-looking firms need to forecast future citation dynamics when making innovation decisions. We focus on a stationary equilibrium with balanced growth to ease the analysis.

Patent quality and citation value are reduced-form modeling devices. Patent quality is a scalar with multiple roles. It governs product-market profitability, which gives rise to quality-directed patent citations, and a new patent's quality step size affects its citation behavior. Quality depreciation captures knowledge obsolescence without the need to model patent and product turnover. Citation value is in the simplest form as an exogenous benefit per new citation received. This model does not attempt to explain such value flows, which represent pecuniary and non-pecuniary benefits the firm gains from others' subsequent uses of its knowledge. They may have property-right origins, such as patent licensing fees or royalty payments, the possibility to receive infringement settlements, signaling values and new customers of the associated products, or they may be a firm's or its R&D department's reputation gains internalized as part of the firm's value. Citation flows are a proxy of knowledge flows and the rate at which firms get such benefits.

The citation process describes the matching between a new patent with a set of existing ones; it is stochastic but governed by endogenous firm decisions. Given firm decisions, the process modifies the standard node copying local mechanism of statistical network formation to accommodate continuous time, a continuum of nodes, and an imperfect selection of input patents' quality.[1] In a nutshell, a new patent may select input patents independently and copy the citation list of a particular input patent imperfectly with mutations; a greater quality step size means fewer copies and more independent citations.[2] A correct copy means to cite the same patent on the list. Mutation occurs with an exogenous probability; a mutated copy means the opportunity to select another patent with similar knowledge content, e.g., within the same technology category, instead of citing the same patent on the list. All cited patents except for correct copies contribute to the new patent's input quality, so they are chosen based on quality. However, quality selections are imperfect with the choice of the highest-quality input patent subject to random shifters; such imperfection ensures tractability and the property that higher-quality patents are more likely to be cited.

Imperfect node copying captures how previously established citation edges influence the formation of new ones; correct copies and mutations are both crucial. Cor-

[1]A textbook version of the node copying model is in section 13.5 of Newman (2018). A local mechanism specifies how each node in a network forms edges with others; various local mechanisms of network formation may lead to common global network characteristics, such as a particular stationary degree distribution. A better-known example is preferential attachment. The term technically describes a specific local mechanism but is often used to represent a class of statistical network formation models that produce "rich-get-richer" dynamics and asymptotic power-law degree distributions. Other network statistics may differ. See, e.g., Newman (2018) and Jackson (2010).

[2]The idea of copying the reference list comes from Price (1976) on paper citations.

rect copies reflect patent-level path dependence, which generates citation transitivity and the "citations-beget-citations" dynamics. Mutation plays two roles. One is to reflect knowledge similarity and spillovers within a research field, thus corresponding to coarser-level path-dependence. The other is that mutated copies allow a new patent to select high-quality inputs at no cost, so the mutation probability reflects the degree to which existing citations help to reduce innovation costs. The citation formation process also captures other realistic aspects of innovations, such as knowledge nonrivalry and innovation uncertainty. Consequently, it produces plausible citation accumulation dynamics and citation-network characteristics. For example, patents get cited more likely at a given time if they have higher own quality levels or have already attracted more and higher-quality citing patents. Moreover, when the correct copying rate is sufficiently high, the stationary citation distribution has an approximate power-law right tail, with the tail exponent determined in equilibrium.

Citation edges correspond to two kinds of externalities or spillovers in the model — knowledge externalities and value externalities. A conventional knowledge *level* externality exists because each new patent builds on the quality level of older ones. The novel knowledge *path* externality occurs when previous citations help to improve new-patent quality, which is captured by a positive mutation probability in the citation process. Both knowledge *level* and *path* externalities involve newer patents exploiting older ones, so the direction is from the past to the future. When the citation value is positive, citation dynamics also lead to three value externalities in the opposite direction, i.e., from future innovations to the past. One is *congestion* resulting from an "excess return" to new-patent quality because every citation it attracts not only brings a contemporaneous citation value but also helps attracts future citations. Congestion is the equilibrium consequence that the excess return incentivizes all firms to choose high new-patent quality levels, so no patents are considerably above the market average. Another is an *abundance* externality such that every new patent means new citation values to distribute among existing ones, which is not the new patent's owner firm's concern. Lastly, a *value sharing externality* is present because a firm cannot claim the full value of a patent by owning it. The reason is that a patent brings a profit stream and a citation-value stream to its owner, and it also helps all cited patents attract further citation values to their respective owners.

It is natural to consider innovation policies through the lens of this model, which suggests additional challenges faced by policymakers and a possible new tool. A critical challenge stems from knowledge externalities, especially with a path-dependent innovation process. It is likely a nature of knowledge creation that hardly responds to any policy measures. The good news is that we may be able to fight (knowledge) externalities with (value) externalities. The model implies that citations could be a new margin for tax-and-subsidy policies. However, the efficacy of such policies hinges on the extent of path externality.

Layout. After the upcoming literature review, the remainder of the article proceeds as follows. Section 2 contains data description and empirical findings. We present the theoretical model in the following order. Section 3 describes the model environment except for details on citation formation. We analyze firms' individual decisions and characterize the equilibrium in Section 4, treating the citation dynamics as a "black box" summarized by simple functions. Then, Section 5 unpacks the network formation process and discusses the network structure in equilibrium. Section 6 completes the analysis. Section 7 contains a brief discussion of calibration and a quantitative illustration of policies. Section 8 concludes. The appendices contain empirical and technical details.

1.1 Related literature

Our empirical findings suggest pathways as another measure of patent quality; our firm-level variables capture each firm's multiple roles in the network. The theoretical framework offers a simple way to incorporate the dynamics of large and expanding networks as part of a general equilibrium with forward-looking agents. This article builds on and contributes to several strands of literature on innovation and growth.

Citation dynamics aside, the model stands close to existing studies on endogenous growth with heterogeneous firms. Each firm's individual problem resembles those in typical Schumpeterian models such as Klette and Kortum (2004) and Lentz and Mortensen (2008) that build on earlier quality-ladder insights by Grossman and Helpman (1991) and Aghion and Howitt (1992), among others. This model also has variety expansion, first put into the equilibrium growth context by Romer (1986, 1990). Innovation quantity and quality choices resemble the intensive and extensive search by Jovanovic and Rob (1990). Aghion, Akcigit, and Howitt (2015) review the literature on endogenous growth. A few more recent contributions are Acemoglu et al. (2018), Olmstead-Rumsey (2020), and Garcia-Macia, Hsieh, and Klenow (2019), for example. Prior growth models that explicitly consider patent citations are scarce. Notable exceptions include Caballero and Jaffe (1993), Eeckhout and Jovanovic (2002), and Akcigit and Kerr (2018). In these cases, citation dynamics are irrelevant to agent payoffs and incentives.

In macroeconomics, research on networks and their formation mainly focuses on the production network. A few study endogenous formation of production and trade networks, such as Grassi (2018), Lim (2017), Oberfield (2018), Acemoglu and Azar (2020), Taschereau-Dumouchel (2020), Ghassibe (2020), Dhyne et al. (2020), and Demir et al.

(2021). This model adopts a different and more tractable approach by considering a hybrid of endogenous and statistical network formation. It complements the literature, especially Oberfield (2018), by endogenizing a different aspect of the network formation process. Here, agent decisions determine the network expansion rate and matching rules of nodes, but the matching outcome is uncertain due to imperfect selections, whereas agents in Oberfield's model decide the matching outcome, given terms of trade and randomly assigned trading partners. Given agent decisions, the statistical process adopts and alters the one by Price (1976) on paper citations, sometimes referred to as the preferential attachment mechanism that became familiar to economists largely because of Jackson and Rogers (2007). Other studies such as Atalay et al. (2011), Atalay (2013), Chaney (2014), and Carvalho and Voigtländer (2014) use similar processes. This model considers node heterogeneity and equilibrium degree dynamics as a step forward in this direction.

The externality analysis contributes to a vast literature on market inefficiency in innovation and R&D policies. Bryan and Williams (2021) give a recent summary. The knowledge path and value externalities induced by citations make prior theories on innovation directions particularly relevant. The majority are purely theoretical. Two dynamic models highlight path-dependent knowledge creation in innovation races. Bryan and Lemus (2017) study state-dependent innovation races in multiple directions. Racing distortion and under-appropriation coexist in the choice of innovation direction. In contrast, a force akin to *over*-appropriation may arise in our model when patent owners benefit from subsequent innovations. Hopenhayn and Squintani (2021) consider horizontal heterogeneity in research returns. Dynamic congestion externalities exist such that high-return problems use too much resources, resulting in excessive and costly researcher reallocation. Our model also has congestion in equilibrium, but it may be an offsetting force to knowledge externality and improve welfare. Akcigit, Hanley, and Serrano-Velarde (2020) and Liu and Ma (2021) emphasize quantitative applications in the context of endogenous growth. Asymmetric externalities due to knowledge spillovers are the source of market inefficiency in both models, leading to asymmetric policy remedies. Akcigit, Hanley, and Serrano-Velarde (2020) distinguish basic and applied research, such that basic research generates larger spillovers than applied research in scale and scope. Liu and Ma (2021) consider sectoral heterogeneity such that further upstream sectors in the innovation network generate larger knowledge spillovers to downstream sectors. Our model shares the concern over knowledge spillovers and their direction with additional emphasis on citation dynamics and value externalities.

The model sheds light on the sector-level knowledge network and cross-sector knowledge diffusion. Section 5.3 shows the aggregated sector-level network structure in equilibrium, with a "home bias" of within-sector citations and quality-driven cross-sector citations. Previous literature acknowledges the role of cross-sector knowledge flows in long-term growth and trade, such as Griliches (1957), Helpman (1998), Jovanovic and Rousseau (2005) Akcigit, Celik, and Greenwood (2016), Cai and Li (2019), Huang et al. (2018a) and Huang et al. (2018b) Cai, Li, and Santacreu (2022). These studies take the sector-level knowledge network as given; here, it is an equilibrium consequence.

The empirical analysis contributes to the literature on firms' market value and patenting activities. See Pakes and Schankerman (1984), Austin (1993), Hall, Jaffe, and Trajtenberg (2001, 2005), Nicholas (2008) and Kogan et al. (2017). These papers try to measure the economic value of patents. Some use citations received as an *ex post* measure of patent quality. Kogan et al. (2017) use the stock market reaction after the patent announcement as a means to construct an appropriate measure of patent quality to study within- as well as between-industry reallocation and growth dynamics. They find positive knowledge spillovers and strong, negative creative destruction to peers as in Bloom, Schankerman, and Van Reenen (2013). Our focus on pathways limits the negative impact of creative destruction and highlights knowledge spillovers.

Our empirical findings also adds to existing patent quality measures summarized by Higham, De Rassenfosse, and Jaffe (2021), from Trajtenberg, Henderson, and Jaffe (1997) to Higham et al. (2019) and Marx and Fuegi (2020). Our definition of a pathfinder patent is similar to novel word in Balsmeier et al. (2018) and the interdisciplinary and cutting-edge measures in Higham, De Rassenfosse, and Jaffe (2021).

2 Empirics

This section describe the primary data sources for the empirical analysis, and it presents the findings on citation network dynamics and their impacts on firm activities. Robustness checks are in Appendix A. Additional tables are in Appendix F.

2.1 Data

We use the 2015 version of the U.S. Patent and Trademark Office (USPTO) Patent Database. It contains detailed information on 5.9 million *utility* patents granted in the U.S. between the years 1976 and 2015, to construct patent level and sectoral level knowledge networks using patent citations. The data also contain each patent's U.S. patent class code to identify its technology category.

We utilize Compustat between 1974 and 2015, retrieved from Wharton Research Data Services to study patenting firms' performances, including their market values. Compustat and the NBER-PDP database are connected using the matching procedure provided in the PDP data, also used by Kogan et al. (2017). Firm-level total factor productivity (TFP) data comes from İmrohoroğlu and Tüzel (2014).

2.2 Pathways, pathfinders, and path-receivers

Patents and citation edges are an observable proxy of knowledge stock and flows. We begin by examining how *transitive* citation edges are at the patent level. Specifically, if patent b cites patent a, and patent c cites b, we compute the empirical probability that c also cites a. The measure is a *directed clustering coefficient* (dcc) modified from a standard global clustering coefficient to account for the directions of knowledge flows.

Fact 0. The patent-citation network is highly transitive, with the dcc averaged at 0.31 over the sample period.

A high transitivity measure suggests highly path-dependent knowledge flows. If patents were to form citations randomly and independently, the citation network's dcc should be $\frac{d}{n-1}$, where d is the average in- and out-degree, and n is the number of patents in the network. The measured dcc is thus magnitudes larger than that of a random network.

We face a few challenges when examining patent-citation dynamics' impacts on patenting firms. First, an innovating firm can simultaneously be a citing firm and a cited one. We need to distinguish these roles for each firm in the data and separately examine any resulting spillovers. Second, new patents bring citation flows but also creative destruction effects. Older knowledge becomes obsolete and replaced; the corresponding product line loses its market share to newer products. Therefore, although it is reasonable to conjecture that citations benefit the cited firms by signaling high quality, observed variables may be under the adverse effects of creative destruction or business stealing.

Our empirical strategy is to focus on cross-category *pathways*. A citation edge is a pathway from category i to another category j if it is the first citation ever made from i to j.[3] By definition, pathways are highly independent of previous citations. A pathway connects two patents. We call the citing one in i a *pathfinder* patent and its owner a pathfinder firm; we call the cited one a *path-receiver* patent and its owner the path-receiver firm. Each pathway's creation is simultaneously a patent-, firm-, and category-level event. Pathfinder and receiver patents belong to different categories, and hence their owner firms are unlikely to be direct competitors at the formation of pathways. Focusing on pathways thus limits the creative destruction effect on the cited firms (path-receivers).

[3] We are aware that the baseline definition of a pathway suffers a left-truncation problem; the first appearance of an i to j citation in the sample may not be the first one ever in history. Therefore, we also consider a few alternative definitions of pathways. See Appendix A for details.

Table 1: Average quality measures of all patents, pathfinders, and path-receivers

	Patent groups		
	All	Pathfinders	Path-receivers
Ex ante quality measures			
Originality †	0.52	0.79	0.59
Number of CPC technology class memberships − 1	0.94	1.10	1.01
Backward citations' pedigree	1.60	1.38	1.48
Average age of backward citations	11.81	9.90	9.12
Grant lag in years	2.63	2.51	2.15
Number of backward citations to scientific literature (front page)†	14.86	18.82	3.96
Number of novel words per patent	0.41	0.29	0.21
Ex post quality measures			
Generality †	0.55	0.64	0.72
Number of forward citations received in 10 years†	3.94	5.51	7.04

† Originality and generality measures are from the NBER-PDP (https://sites.google.com/site/patentdataproject/), calculated using 1976–2006 USPTO patent citation data. When calculating "Number of forward citations in 10 years", we drop the last 10 years' data to get rid of the truncation concern. Citation to scientific literature data are from Marx and Fuegi (2020).

Fact 1. Pathways indicate high patent quality. Compared with an average patent, pathfinder and path-receiver patents attract more forward citations, have higher generality and originality scores and wider CPC technology class memberships, and cite younger patents.

Fact 1 says that pathways tend to appear between high-quality patents. We draw from the literature a set of commonly used quality measures, such as originality, generality (Trajtenberg, Henderson, and Jaffe, 1997), number of forward citations, number of Collaborated Patent Classification (CPC) technology class memberships, and several others used in Higham, De Rassenfosse, and Jaffe (2021).[4] Table 1 reports the results.

Cross-category pathways as an easily observable indicator of patent quality have the following features. First, a pathway is an *edge*-based indicator by definition, highlighting citation dynamics. Existing quality measures focus on individual patent (*node*) characteristics. Second, a pathway is simultaneously an *ex ante* quality measure and an *ex post* one. To the pathfinder (*citing*), it is *ex ante* and relates to other ex ante measures such as "originality". To the path-receiver (*cited*), a pathway happens *ex post* and adds to other *ex post* measures such as "generality" and forward citation counts. The quality-driven citation process in Section 5 explicitly captures these observations.

Pathfinder patents have above-average quality by most quality measures in the literature, except for backward citations' pedigree, grant lag, and novel word count. Higham, De Rassenfosse, and Jaffe (2021) suggest lower backward citations' pedigree and shorter grant lag predict lower patent quality. Balsmeier et al. (2018) use weighted novel word

[4]Backward citations' pedigree measure is granted from Higham et al. (2019). We thank Kyle Higham for sharing the pedigree data with us.

Table 2: Innovation scale and scope of pathfinder and path-receiver firms

	All patenting firms	Pathfinder firms	Path-receiver firms
Mean patent stock	45.72	250.86	451.30
Mean patent categories	1.84	5.99	7.75

count to evaluate a patent's originality. Pathfinders are likely *new combinations* of existing ideas, such as finding a new use of an existing product. Hence, they need not cite other popular patents (backward citations' pedigree) or create a brand new idea (novel word). Consequently, pathfinders' novelty is likely evident to judge, so they are quickly granted.

Path-receiver patents have similar quality measures as pathfinder patents. An exception is the number of backward citations to scientific literature, measuring a patent's relation to frontier scientific research. A lower score in this measure means path-receivers are relatively well-established and readily applicable compared to average patents. Intuitively, when inventors utilize knowledge from a new sector for the first time, they start from the best-known and tested patents instead of fancy and frontier ones.

Which firms own pathfinder and path-receiver patents? Table 2 shows that that pathfinder and path-receiver firms are several times larger than an average patenting firm in terms of patent number and scope. We caution that the observation does not imply that larger firms' patents have higher quality on average; Arora et al. (2023) find a negative relationship between firm size and the average patent quality measured by forward citation count. A possible explanation is that firms need to accumulate enough knowledge in either the citing or the cited category before creating a pathway between two previously disconnected categories. Larger firms are more likely to satisfy this condition.

2.3 Impacts of citation network dynamics on firms

This section examines the impact of network dynamics on firms' innovation and product-market performance. The first task is to construct measures of each firm's four kinds of involvement in pathway creations. A pathway from category i to category j directly connects the pathfinder and receiver firms. It also establishes that knowledge in j is useful in producing new knowledge (patents) in i. Therefore, such an event may have spillover effects on other firms (peers) in i and j. We are interested in both the direct impacts on pathfinder and receiver firms and the spillover impacts to their peers.

In each sample year t, we calculate each firm f's patent-stock share weighted counts of pathways to capture f's four possible roles: weighted number of pathways received and found by firm f, $\text{wpr}_{f,t} = \sum_j w^f_{j,t} pr^f_{j,t}$ and $\text{wpf}_{f,t} = \sum_j w^f_{j,t} pf^f_{j,t}$; weighted number of pathways received and found by f's peers, $\text{wpr}^{peer}_{f,t} = \sum_j w^f_{j,t} pr^{-f}_{j,t}$ and $\text{wpf}^{peer}_{f,t} =$

$\sum_j w^f_{j,t} pf^{-f}_{j,t}$, where $pr^f_{j,t}$, $pf^f_{j,t}$, $pr^{-f}_{j,t}$ and $pf^{-f}_{j,t}$ are numbers of pathways received and found by firm f itself and other firms $-f$ in year t category j, respectively. The weights are f's patent stock portfolio across categories, such that $w^f_{j,t} = N^f_{j,t}/N^f_t$, where $N^f_{j,t}$ is firm f's patent stock in category j at t, and $N^f_t = \sum_k N^f_{k,t}$ is its total patent stock at t. The weights reflect a firm's various exposures to pathway creations: a pathway from or to category j likely has greater impacts on firms with larger patent shares in j.

The four weighted pathway counts measure each firm's direct and indirect involvement in pathway creations in year t. We look at the changes in firm activities over a horizon τ of one to five years after pathways' appearance. The regression's general form is

$$\log\left(X^f_{t+\tau}/X^f_t\right) = \beta_{pr,\tau}\mathrm{lwpr}_{f,t} + \beta^{peer}_{pr,\tau}\mathrm{lwpr}^{peer}_{f,t} + \beta_{pf,\tau}\mathrm{lwpf}_{f,t} + \beta^{peer}_{pf,\tau}\mathrm{lwpf}^{peer}_{f,t} + \boldsymbol{Z}_{f,t} + \mathrm{err}_{f,t,\tau}, \tag{1}$$

where X^f_t is any variable of interest, such as sales, for each firm f in each year t, and we take the logarithms of the four weighted pathway counts and then standardize them.[5] The term $\boldsymbol{Z}_{f,t}$ represents a set of controls including firm and year fixed effects, together with other variables that differ across regressions.

Citation dynamics and innovation activities. Let X^f_t be firm f's patent stock at t. The dependent variable in eq. (1) is the firm's innovation rate over a five-year horizon. Controls in $\boldsymbol{Z}_{f,t}$ include firm f's patent stock at t and the number of technology categories in which firm f has patent application at t.

Fact 2. Over a five-year horizon, after a pathway appears, a patenting firm is more likely to expand into a new technology category if the firm is directly or indirectly involved in pathway creation.

Table 3 shows the results of regression on firm-level innovation rates. Pathways indicate high qualities of both the finder and the receiver patents. As expected, the innovation rates of path-receiver firms see increases afterward, reflected by the positive and significant coefficients of weighted pathways received in logarithms. Furthermore, path-receiver firms' peers in the same category also show higher innovation rates, suggesting positive spillovers. Additionally, we want to emphasize that our empirical results still hold when controlling firm fixed effects other than industry and time fixed effects, while in the literature, most empirical studies only control industry and time fixed effects.

We can split firm innovation rates into extensive- and intensive-margin growth. Extensive-margin growth captures a firm's patent application growth in new technology categories

[5] Specifically, we let $\mathrm{lwpr}_{f,t} = \log(\mathrm{wpr}_{f,t} + 1)$, treating the rest in the same way, and then standardize them so that we can compare the coefficients.

Table 3: Firms' cumulative innovation rates up to five years after pathways appear: regression results of eq. (1) with X_t^f being firm f's patent stock at t.

	Horizon τ in years				
Each firm's involvement as	1	2	3	4	5
a pathfinder (lwpf)	-0.00137	-0.00264	-0.00525**	-0.00692***	-0.00619*
	(-0.92)	(-1.64)	(-3.00)	(-3.34)	(-2.15)
a peer to pathfinders (lwpf_peer)	0.0662***	0.0610***	0.0507***	0.0439***	0.0406***
	(20.32)	(16.20)	(11.20)	(8.56)	(7.20)
a path-receiver (lwpr)	0.00170*	0.00360***	0.00399***	0.00368***	0.00324**
	(2.36)	(4.18)	(4.08)	(3.40)	(2.82)
a peer to path-receivers (lwpr_peer)	0.107***	0.0859***	0.0616***	0.0462***	0.0368***
	(33.20)	(23.62)	(14.23)	(9.37)	(6.77)
N	921983	753471	610015	500379	415652
R^2	0.409	0.465	0.457	0.436	0.425

Controls in $\boldsymbol{Z}_{f,t}$ include firm patent stock and the number of patenting categories, both in logarithms, and firm and year fixed effects. Observations are clustered by firm. All variables are standardized. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

that this firm has never patented before; intensive-margin growth measures a firm's patent application growth in existing categories. In robustness checks, we run separated regressions on extensive and intensive innovation rates, respectively, and find that firms benefit from pathways through extensive growth rather than intensive growth.

The coefficients capturing the spillover effects on peers in Table 3 appear to be magnitudes larger than coefficients of direct impacts. The reasons are twofold. One is that, as Table 2 shows, pathfinder and receiver firms tend to be much larger than their peers and thus less sensitive on the growth margin. In robustness check regressions with top 5 percent large firms only, the spillover effects on peers are much smaller, supporting this conjecture. The other is that pathways are relatively rare events, and most firms are never directly involved, resulting in small variations in corresponding measures. In contrast, firms are much more likely to be indirectly involved, and the variations in corresponding measures are magnitudes larger.

Citation dynamics and firm performances. We turn to firms' product-market performance measures and focus on publicly traded firms only. In this case, X_t^f in eq. (1) can be a firm's profit, sales, market value, total employment, capital stock, or productivity. Controls in $\boldsymbol{Z}_{f,t}$ now include two measures of a firm's innovation output in logarithms following Kogan et al. (2017), where one is a stock-market based measure of patents' dollar values, and the other is citation-weighted patent stock. Additionally, we include several size controls such as log-scaled employment, capital, and X_t^f if applicable. Firm and year fixed effects remain in place.

Fact 3. Public pathfinder and receiver firms grow more in sales, profit, employment, and

Table 4: Firms' cumulative growth rates in profits and employment up to five years after pathways appear: regression results of eq. (1) with X_t^f being firm f's profit or number of employees at t.

	Horizon τ in years				
Each firm's involvement as	1	2	3	4	5
Panel A. Profit growths					
a pathfinder (lwpf)	0.00851***	0.0139***	0.0142***	0.0146***	0.0119**
	(4.37)	(5.33)	(4.84)	(4.51)	(3.29)
a peer to pathfinders (lwpf_peer)	0.00722	0.00227	-0.00470	0.00606	-0.00244
	(1.78)	(0.40)	(-0.81)	(0.97)	(-0.39)
a path-receiver (lwpr)	0.0159***	0.0194***	0.0162***	0.0179***	0.0213***
	(9.08)	(7.83)	(5.55)	(5.55)	(5.70)
a peer to path-receivers (lwpr_peer)	0.00958*	0.0115	0.0157**	0.0103	0.0161**
	(2.13)	(1.89)	(2.80)	(1.77)	(2.73)
N	32699	29860	27231	24909	22781
R^2	0.387	0.472	0.536	0.588	0.633
Panel B. Employment growths					
a pathfinder (lwpf)	0.00874***	0.0127***	0.0112***	0.00960***	0.00905**
	(6.37)	(6.49)	(4.97)	(3.55)	(3.03)
a peer to pathfinders (lwpf_peer)	0.00995***	0.0117**	0.00950*	0.0101*	0.00308
	(4.17)	(3.17)	(2.11)	(1.99)	(0.54)
a path-receiver (lwpr)	0.00986***	0.0137***	0.0148***	0.0173***	0.0181***
	(7.78)	(7.12)	(6.15)	(6.27)	(5.68)
a peer to path-receivers (lwpr_peer)	0.0112***	0.0157***	0.0159***	0.0134**	0.0160**
	(3.87)	(3.98)	(3.52)	(2.76)	(3.23)
N	34599	31437	28541	25987	23681
R^2	0.339	0.446	0.523	0.583	0.627

Controls in $\boldsymbol{Z}_{f,t}$ include a firm's stock-market based measure of patents' dollar values and its citation-weighted patent stock following Kogan et al. (2017), both in logarithms, size controls, and firm and year fixed effects. Observations are clustered by firm. All variables are standardized. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

capital up to five years after pathways appear. Spillover effects to their peer firms in the same direction exist but can be weak.

Table 4 only presents the regression results on profit and employment growths to be concise; the rest are in the appendix. Consistent with intuitions and previous studies, the results suggest that a pathfinder firm directly benefits from owning a new, high-quality pathway patent, even after controlling for the firm's innovation output, size, and a fixed effect. These firms tend to have higher sales revenue and profits, show higher growths in capital and employment, and have higher market values. Peer firms innovating in the same category as the pathfinder patent also gain from the discovery of a new knowledge pathway, because they can copy the new route to find wider source of knowledge input in future innovation. Relatively speaking, spillovers from a pathfinder to its peers are relatively weak. A potential explanation is that the positive effect of technological spillovers is offset by the negative one of business stealing brought by the pathfinder patent.

A path-receiver firm benefits from receiving such a citation, with coefficient values comparable to a pathfinder's or even more prominent. With the set of controls in place, a firm's (revenue-based) productivity increases after pathway appearance only if the firm is the path-receiver. Spillovers from a path-receiver to its peers are positive and can be significant. By definition, pathfinder patents have limited creative destruction effects on those in the receiving category. In contrast, if knowledge is somewhat substitutable in the pathway receiving category, then the newly discovered use of the path-receiver patent brings attention and potential new uses of all patents in the category, benefiting peer firms in the used category through forward citations, royalty payment and reputation.

3 Model

The model studies firms' innovation incentives and the aggregate implications in a general equilibrium framework. Knowledge exists as patents. New patents use existing ones as inputs. We use citation dynamics to capture path-dependent knowledge creation and related spillovers. In particular, existing citations describe previous knowledge flows, and they influence how new patents make citations. Furthermore, citation is quality-directed such that new patents are more likely to use (cite) higher-quality existing patents as inputs.

Time is continuous with an infinite horizon, $t \in [0, \infty)$. The economy has a unit-mass continuum of identical price-taking households who work and consume. The economy has a fixed continuum $[0, F]$ of risk-neutral firms, $F > 0$, indexed by $f \in [0, F]$. Firms hire, produce, and innovate. We do not consider firm entry or exit. Each firm f is endowed with patent stock $n^f(0) \in \mathbb{N}$ at time 0, and accumulates patents over time. Newer patents use old ones as inputs and cite them. A patent corresponds to a product and its production technology, and it belongs to its inventing firm. Patent ownership gives a firm the right to produce, price, and sell the corresponding product. The model abstracts from patent tradings. The growing set $\mathcal{N}(t) = [0, N(t)]$ includes all available products and patents at any time t, where $N(t)$ is the aggregate knowledge stock. Firms and patents belong in J sectors, indexed by $j \in \{1, \ldots, J\}$. Each firm innovates in one sector, and every sector has a positive measure of innovating firms. To keep the model transparent, we assume symmetric sectors without any real heterogeneity. The only role of sectoral classification is to guide the citation dynamics. Product markets are monopolistically competitive, and firms set the product prices. Labor markets are perfectly competitive.

3.1 Household preferences

A unit-measure continuum of identical household populates the economy; they discount the future at rate $\rho > 0$. A representative household fully summarizes their behavior, with the lifetime preferences described by $U = \int_0^\infty \exp(-\rho t) \left(\log C(t) - \chi \frac{R(t)^{\sigma+1}}{\sigma+1} \right) \mathrm{d}t$, where $R(t)$ is the research labor supplied with the elasticity inverse $\sigma > 0$, and $C(t)$ is the consumption compound such that it aggregates all available products at time t with constant elasticity of substitution (CES), given as $C(t) = \big(\int_0^{N(t)} z(\omega, t)^{\frac{1}{\nu}} c(\omega, t)^{\frac{\nu-1}{\nu}} \mathrm{d}\omega \big)^{\nu/(\nu-1)}$, where $\nu > 1$ is the substitution elasticity among products, $c(\omega, t)$ is the consumption quantity of product ω at t, and the product-specific taste shifter $z(\omega, t)$ is the *quality* of ω at t.[6] Households take each $z(\omega, t)$ as given; quality depreciates at a fixed rate $\delta > 0$ such that $\dot{z}(\omega, t) = -\delta z(\omega, t)$, $\forall \omega \in \mathcal{N}(t)$. Households own all firms. Household income flow consists of labor income and profit flow from firms. Households supply production labor $L > 0$ inelastically at any time. Henceforth, we choose the production labor as the numeraire. Production labor represents all factor endowments that are used in production. Denote the wage rate of research labor as $w(t)$. The representative household's budget constraint at any t reduces to $\int_0^{N(t)} p(\omega, t) c(\omega, t) \mathrm{d}\omega = L + w(t) R(t) + \Pi(t)$, where $\Pi(t)$ is the aggregated flow profits. Standard cost minimization yields the demand function of each product variety such that

$$c(\omega, t) = z(\omega, t) C(t) \left(\frac{p(\omega, t)}{P(t)} \right)^{-\nu}, \tag{2}$$

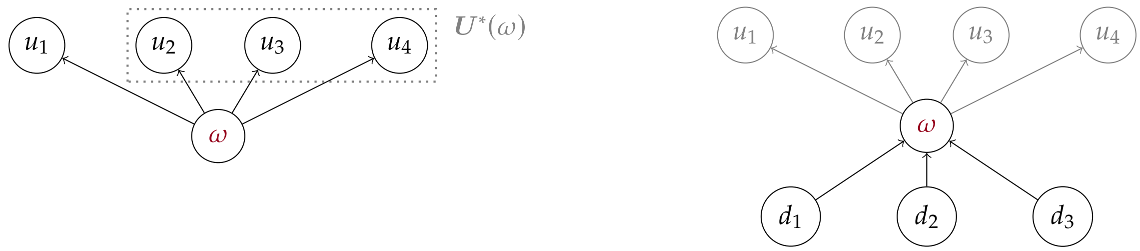
where $P(t)$ is the ideal price index, given as $P(t) = \big(\int_0^{N(t)} z(\omega, t) p(\omega, t)^{-(\nu-1)} \mathrm{d}\omega \big)^{-1/(\nu-1)}$. The optimal supply of of research labor is $R(t) = \big(w(t) / [\chi P(t) C(t)] \big)^{1/\sigma}$.

3.2 Firms

Firms are risk-neutral and forward looking. The numeraire choice implies that firms discount future profits at the same rate ρ as the household. Consider a typical firm f with $n^f(t) = n \in \mathbb{N}$ patents at time t, indexed by $\omega = 1, \ldots, n$. The aggregate knowledge stock is $N(t)$. The market wage rate for research labor is $w(t)$.

Innovation rate. Firm f gets new ideas according to a Poisson process, and f sets the arrival rate. Quality of existing patents affects this choice. Suppose firm f decides to use $\widehat{R}_x > 0$ units of research labor, and its existing patent $\omega = 1, \ldots, n$ has a time-t quality level z^ω. A constant-returns-to-scale (CRS) function $I : \mathbb{R}_+^2 \to \mathbb{R}_+$ determines the time-t innovation rate $\widehat{I}$ as $\widehat{I} = I\left(N(t) \widehat{R}_x,\ \sum_{\omega=1}^n z^\omega \right) = \left(\sum_{\omega=1}^n z^\omega \right) \times I\left(\frac{N(t) \widehat{R}_x}{\sum_{\omega=1}^n z^\omega},\ 1 \right)$. $I(\cdot)$ strictly

[6] Appendix D discusses households' degree of love-of-variety, which does not affect firm problems in the market equilibrium.



(a) Patent ω cites $\boldsymbol{U}(\omega) = \{u_1, u_2, u_3, u_4\}$ at birth; its input quality depends on $\boldsymbol{U}^*(\omega) = \{u_2, u_3, u_4\}$.

(b) At some given time t, patent ω has attracted $\boldsymbol{D}(\omega, t) = \{d_1, d_2, d_3\}$.

Figure 1: An illustration of patent ω's local network structure: fixed $\boldsymbol{U}(\omega)$ and expanding $\boldsymbol{D}(\omega, t)$ since birth. Each directed citation edge goes from the *citing* patent to the *cited* one.

increases in both arguments. That is, from t to $t + \mathrm{d}t$, $\mathrm{d}t \to 0$, firm f gets a new idea with approximate probability $\widehat{I}\mathrm{d}t$. Using the same amount of innovation resources, firms with more patents of higher quality levels get new ideas more frequently. Equivalently, we use a univariate cost function $c_x : \mathbb{R}_+ \to \mathbb{R}_+$ to get the research labor units $\widehat{R}_x$ required to achieve a chosen innovation rate $\widehat{I} = \sum_{\omega=1}^{n} z^{\omega} \widehat{x}$ at t, such that $\widehat{R}_x = \frac{1}{N(t)} \sum_{\omega=1}^{n} z^{\omega} \cdot c_x\left(\widehat{x}\right)$, where $\widehat{x} \geq 0$ may be referred to as the innovation intensity, and $c_x(\cdot)$ is at least twice differentiable with $c_x(0) = 0$, $c_x' > 0$, and $c_x'' > 0$.

Citations and patent quality. Conditional on a new idea's arrival at t, firm f turns it into a new patent and decides on its quality. Each new patent must cite c existing ones as inputs, where $c \in \mathbb{N}$ is a fixed number and $c > 1$.[7] Each of the c citations goes to an existing patent owned by a different firm. As the initiator of citation edges, a citer firm at time t faces a per edge cost of $\frac{1}{N(t)}\phi$ in numeraire units, with $\phi \geq 0$. Therefore, a time-t new patent citing c existing ones requires a fixed cost of $\frac{1}{N(t)}\phi c$.

Firm f sets the new patent's quality level by choosing a "step size" of quality improvement. Quality step size serves two roles. One is the size of quality increment over a baseline input quality. The more interesting role is that the quality step size also governs said input quality. The idea is that the size of technological advancement also reflects the degree to which the new patent absorbs and exploits knowledge from existing patents.

To achieve a new patent's quality "step size" $\widehat{q}$ at t, firm f must hire $\widehat{R}_q$ units of research labor, such that $\widehat{R}_q = \frac{1}{N(t)} c_q(\widehat{q})$, where the cost function $c_q : [\underline{q}, 1] \to \mathbb{R}_+$ for some $\underline{q} \in (0, 1)$ is at least twice differentiable, with $c_q(\underline{q}) = 0$ and $c_q' > 0$, $c_q'' > 0$ over $(\underline{q}, 1)$.

Let $\boldsymbol{U}(\omega)$ denote the realized set of (upstream) input patents cited by a new patent ω, which has a fixed cardinality $|\boldsymbol{U}| = c$.[8] Only a subset of the cited patents $\boldsymbol{U}^* \subseteq \boldsymbol{U}$

[7] The restrictions that c is an integer and that c is common to all new patents are for easy interpretations but not required, thanks to mean-field approximation.

[8] When not causing confusion, we drop patent identity ω and time t to ease notation.

is relevant for the new patent's quality, with $1 \le |\boldsymbol{U}^*| \le c$. We postpone the detailed formation process of $\boldsymbol{U}$ and its subset $\boldsymbol{U}^*$ until Section 5. What matters at this stage is that firm f understands the process such that it expects $|\boldsymbol{U}^*|$ to increase in $\widehat{q}$. Given $\boldsymbol{U}^*$, the "input quality" is $\frac{1}{c}\sum_{u\in\boldsymbol{U}^*} z(u,t)$, where $z(u,t)$ is the time-t quality of a cited input patent u. See Figure 1a for an illustration. Input quality thus increases in the size of $\boldsymbol{U}^*$ and the quality level of each patent in $\boldsymbol{U}^*$. Existing patents are optimally selected into $\boldsymbol{U}^*$ subject to some friction, to be specified in Section 5. Consequently, when hiring research labor $\widehat{R}_q$ to achieve quality improvement $\widehat{q}$ at t, firm f anticipates its new patent's quality to be

$$z^{new}(\widehat{q},\boldsymbol{U}^*,t) = \frac{1}{c}\sum_{u\in\boldsymbol{U}^*} z(u,t) + \widehat{q}, \tag{3}$$

which takes into consideration how $\widehat{q}$ affects $\boldsymbol{U}^*$'s formation.

Firms understand knowledge obsolescence reflected by quality depreciation rate δ. Depreciation captures a flavor of creative destruction and patent expiration without the need to model node and edge removal from the patent citation network.

Production and product pricing. At any time t, the owner firm of an existing patent ω can produce the corresponding product using production labor as the only input, according to a simple production technology $y(\omega,t) = l(\omega,t)$, where $l(\omega,t)$ is the production labor input hired at the competitive unit wage rate, and $y(\omega,t)$ is the output level priced under monopolistic competition. Firm f with n patents at t operates as many product lines. Only the owner firm of a patent can produce the product.

Value of citations. Let $\boldsymbol{D}(\omega,t)$ be patent ω's set of existing (downstream) citers at time t, containing all patents that has cited ω by t. Patent ω's *indegree* at t simply counts the number of citations has attracted by t, denoted as $|\boldsymbol{D}(\omega,t)|$. Upon birth, a new patent has zero indegree and an empty citer set $\boldsymbol{D} = \emptyset$; the set expands over time. See Figure 1b for an illustration. Citations received bring additional values to the cited firms. Assume that a firm's gain from receiving a new citation at time t is $\frac{1}{N(t)}\phi$, which coincides with the cost of each citation made at t.

The citation value ϕ is a reduced-form modeling device to introduce the potential benefit associated with receiving citations. The model is agnostic about the origin of ϕ, but it could correspond to a few real-life scenarios. To a cited firm, the citation value ϕ represents any benefit it gains when other firms use its knowledge. For example, patent owner firms may directly receive royalty payments or settlement of infringement disputes when others use their knowledge; they may also indirectly benefit from reputation gains and signaling values associated with citations received. To a citing firm, ϕ captures a form of search cost to find suitable input knowledge, and it also reflects innovation costs other than researcher wage. Note that the flows of ϕ do not directly affect household income or claim resources; the economy remains closed without imposing additional conditions. Consequently, the setup has an advantage that citation dynamics do not lead to hardwired loss or gain in the economy.

Equilibrium concept. We look for a stationary equilibrium with balanced growth assuming an appropriate law of large numbers. The competitive market for production labor always clears at the unit wage rate. A constant wage rate $w(t) = w > 0$ clears the competitive market for research labor at any time with a time-invariant supply R. Firms optimally set product prices under monopolistic competition given the price index $P(t)$ and its evolution. Aggregate knowledge stock $N(t)$ and sectoral knowledge stocks grow at a common and constant rate $g > 0$, such that $\dot{N}(t) = gN(t)$. New patents have a constant mean quality $\overline{z}$ regardless of time of birth. Firms form forecast of future network dynamics when making innovation decisions. Patent-level network dynamics and stationary cross-sectional distribution are consistent with firm decisions and forecast.

4 Analysis and Equilibrium

In general, the state of any patent ω at time t includes its quality $z(\omega, t)$ and its local network structure captured by a fixed input set $\boldsymbol{U}(\omega)$ and an expanding citer set $\boldsymbol{D}(\omega, t)$. Each innovating firm owns an evolving collection of $n^f(t)$ patents with $\boldsymbol{S}^f(t) = \{(z(\omega, t), \boldsymbol{U}(\omega), \boldsymbol{D}(\omega, t))\}_{\omega=1}^{n^f(t)}$. To make innovation decisions, firm f must form forecast of the law of motion of $\boldsymbol{S}^f(t)$, which depends on f's and other firms' decisions.

In this section, we set up firms' problems *assuming* and taking as given simple functions describing the expected citation dynamics in equilibrium. We proceed to analyze firm decisions and formalize the equilibrium definition. As such, we highlight (i) how citations enter a firm's problem and (ii) what needs adjusting for a standard innovation model to accommodate citation-network dynamics. Derivation details are in Appendix B.

4.1 Taken as given: aggregate variables and expected citation dynamics

Each individual firm takes as given aggregate variables in equilibrium, including a constant growth rate $g > 0$ and new patents' constant mean quality $\overline{z} > 0$.. Firms also take as given the endogenous citation dynamics.

Cross-sectional average of patent quality. The cross-sectional average quality of all existing patents is a "market quality" level. Given g, the stationary cross-sectional distribution of patent age $\tau \geq 0$ is exponential with a probability density function (pdf)

$g\exp(-g\tau)$. Therefore, the stationary market quality satisfies

$$\langle z \rangle = \int_0^\infty g\exp(-g\tau)\overline{z}\exp(-\delta\tau)\mathrm{d}\tau = \frac{g}{g+\delta}\overline{z}, \tag{4}$$

where $\langle \cdot \rangle$ is the standard operator for cross-sectional mean in the network.

Expected quality of a new patent. A time-t new patent's quality depends on the chosen quality step size $\widehat{q}$ and the selection of input patents in eq. (3). Given g and $\overline{z}$, the expected new-patent quality is a deterministic and time-invariant function $z^e(\widehat{q};\ g, \overline{z})$: $[\underline{q}, 1] \rightarrow \mathbb{R}_+$ with $z^e(\widehat{q}) = \mathbb{E}[z^{new}(\widehat{q}, \boldsymbol{U}^*) \mid \widehat{q}]$, where $z^e(\widehat{q})$ is short-hand for $z^e(\widehat{q};\ g, \overline{z})$, and the expectation operator takes into account input selection. Let $z^e(\cdot)$ be strictly increasing and differentiable, with $(z^e)' \geq 1$, so an increase in step size $\widehat{q}$ leads to a larger increase in the expected new-patent quality. The exact form of $z^e(\cdot)$ will be given in Proposition 2.

Expected indegree growth. The law of motion of a patent's *indegree* $|\boldsymbol{D}(\omega, t)|$ affects firm value because a new citation attracted at t means a flow benefit of $\frac{1}{N(t)}\phi$. We allow a patent ω's indegree evolution to depend on its quality $z(\omega, t) = z$ and *quality-weighted* indegree $Z_D(\omega, t) = Z_D$, defined as the summation of each citer patent's quality,

$$Z_D(\omega, t) \equiv \sum\nolimits_{d \in \boldsymbol{D}(\omega, t)} z(d, t). \tag{5}$$

To wit, a patent ω attracts citations faster if it has higher quality or has more higher-quality citers. We need two functions to describe each patent's expected indegree evolution. Let $h, H : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ be linear functions of z and Z_D, such that

$$h(z, Z_D) = \lim_{\mathrm{d}t \rightarrow 0^+} \frac{1}{\mathrm{d}t}\left(\mathbb{E}[|\boldsymbol{D}(\omega, t+\mathrm{d}t)| - |\boldsymbol{D}(\omega, t)| \mid z, Z_D]\right) = z \cdot h_z + Z_D \cdot h_D, \tag{6}$$

$$H(z, Z_D) = \lim_{\mathrm{d}t \rightarrow 0^+} \frac{1}{\mathrm{d}t}\left(\mathbb{E}[Z_D(\omega, t+\mathrm{d}t) - Z_D(\omega, t) \mid z, Z_D]\right) = z \cdot H_z + Z_D \cdot H_D, \tag{7}$$

where endogenous coefficients $h_z > 0$, $h_D \geq 0$, $H_z > 0$, and $H_D \geq -\delta$ depend on g and $\overline{z}$, $h(\cdot)$ describes the expected law of motion of indegrees, and $H(\cdot)$ describes that of quality-weighted indegrees. The forms of (h_z, h_D, H_z, H_D) will be given in Proposition 3.

4.2 Optimal product pricing and equilibrium research labor supply

Each firm optimally prices each product facing the product-specific demand curve in eq. (2) and the unit wage rate. This decision is static and separable from the rest of the firm's problem. A standard pricing rule follows, such that every product has the same price $p = \frac{\nu}{\nu-1}$. The ideal price index $P(t)$ satisfies $P(t) = \frac{\nu}{\nu-1}\left(N(t)\langle z \rangle\right)^{-1/(\nu-1)}$. When the market for production labor clears at $L = \int_0^{N(t)} l(\omega, t)\mathrm{d}\omega$, we have the aggregate consumption compound given as $C(t) = L\big(N(t)\langle z\rangle\big)^{1/(\nu-1)}$, which grows at a constant rate $\frac{g}{\nu-1}$. The aggregate flow profit resulting from production and sales is $\frac{L}{\nu-1}$. The supply of research labor can thus be written as

$$R = R^s(w) \equiv w^{\frac{1}{\sigma}} \left(\chi \frac{\nu}{\nu-1} L\right)^{-\frac{1}{\sigma}}. \tag{8}$$

An equilibrium requires $\Pi = \frac{L}{\nu-1} - wR$ in the household's budget. The optimal output of each variety follows. The maximum instantaneous profit generated by each product line depends on its current quality z. A unit measure of patents with identical z at t bring a flow profit of $\widetilde{\pi}(z,t) = \frac{1}{N(t)}\pi z$, where the time-invariant profit coefficient $\pi > 0$ is given as

$$\pi = \frac{1}{\langle z\rangle}\frac{L}{\nu-1}. \tag{9}$$

4.3 Firm value, innovation decisions, and equilibrium definition

This section shows a firm's value function along an equilibrium path given a fixed aggregate growth rate g, a constant mean quality level $\overline{z}$ of new patents chosen by other firms, and a fixed researcher wage rate w. Firms understand the citation dynamics summarized by functions $h(\cdot)$ and $H(\cdot)$ and expect the new-patent quality to be $z^e(\cdot)$.

Consider firm f at time t with $n^f(t) = n$ patents indexed by $\omega = 1, \ldots, n$. Recall that firms discount future value flows at the same rate ρ as the household because of the numeraire choice. Firm f's flow profit depends on each patent's time-t quality $z(\omega,t) = z^\omega, \forall\omega$; its expected indegree dynamics described by $h(\cdot)$ and $H(\cdot)$ depend on each patent's z^ω and quality-weighted indegree $Z_D(\omega,t) = Z_D^\omega$. Note that Z_D represents the quality of patents owned by other firms. A patent's input set affects it initial quality level via the subset $\boldsymbol{U}^*(\omega)$ at birth, but it remains unchanged afterwards. Therefore, the list of existing patents' states reduces to $\boldsymbol{S} = \{(z^\omega, Z_D^\omega)\}_{\omega=1}^n$.

Let $\widetilde{V}_n : \mathbb{R}_+^{2n+1} \to \mathbb{R}$ be the value of firm f with n patents in time-t numeraire units. Along any stationary growth path, firm f's value satisfies the following equation,

$$\begin{aligned}\rho\widetilde{V}_n(t,\boldsymbol{S}) = \max_{\widehat{x},\,\widehat{q},\,\boldsymbol{U}^*} \Bigg\{ & \frac{\pi}{N(t)}\sum_{\omega=1}^n z^\omega + \frac{\phi}{N(t)}\sum_{\omega=1}^n h(z^\omega, Z_D^\omega) + \frac{\mathrm{d}}{\mathrm{d}t}\,\mathbb{E}\left[\widetilde{V}_n(t,\boldsymbol{S})\right] \\ & - w\frac{1}{N(t)}\sum_{\omega=1}^n z^\omega c_x(\widehat{x}) - \sum_{\omega=1}^n z^\omega \widehat{x}\left(w\frac{1}{N(t)}c_q(\widehat{q}) + \frac{\phi}{N(t)}c\right) \\ & + \sum_{\omega=1}^n z^\omega \widehat{x}\Big(\mathbb{E}[\widetilde{V}_{n+1}(t, \boldsymbol{S} \cup_+ (z^{new}(\widehat{q},\boldsymbol{U}^*),0)) \mid \widehat{q}] - \widetilde{V}_n(t,\boldsymbol{S})\Big)\Bigg\},\end{aligned} \tag{10}$$

subject to $\widehat{x} \geq 0$ and $\widehat{q} \in [\underline{q}, 1]$, where "$\cup_+$" in the last line means appending a new element to the list of states, and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\widetilde{V}_n(t, \boldsymbol{S})\right] = \frac{\partial \widetilde{V}_n}{\partial t} + \sum\nolimits_{\omega=1}^{n}\left(-\delta z^{\omega}\frac{\partial \widetilde{V}_n}{\partial z^{\omega}} + H(z^{\omega}, Z_D^{\omega})\frac{\partial \widetilde{V}_n}{\partial Z_D^{\omega}}\right). \tag{11}$$

On the right-hand side of eq. (10), the first item is the total flow profit that firm f receives, and the second one is the expected flow benefit when patents attract new citations at rate $h(\cdot)$ given in eq. (6). The third item, expressed in eq. (11), is value evolution over time. Each patent's quality depreciates at rate δ, and its weighted indegree is expected to change by $H(\cdot)$ as given in eq. (7). The rest of eq. (10) concerns the option value of innovation if the arrival rate of new patents is $\sum_{\omega=1}^{n} z^{\omega}\widehat{x} \geq 0$, and the expected new-patent quality is $z^e(\widehat{q}) = \mathbb{E}[z^{new}(\widehat{q}, \boldsymbol{U}^*) \mid \widehat{q}]$ with optimally selected input patents. New patents enter with zero indegrees, so their initial state is $(z^{new}(\widehat{q}, \boldsymbol{U}^*), 0)$.

We simplify a firm's value function by expressing it relative to the growth trend of the aggregate knowledge stock, i.e., $V_n(\boldsymbol{S}) = N(t)\widetilde{V}_n(t, \boldsymbol{S})$, with $V_n : \mathbb{R}_+^{2n} \to \mathbb{R}$. We guess and verify that firm value summarizes patent values in an additive form, give as $V_n(\boldsymbol{S}) = \sum_{\omega=1}^{n} v(z^{\omega}, Z_D^{\omega})$, where the growth-adjusted per patent value function $v : \mathbb{R}_+^2 \to \mathbb{R}$ is linear, such that

$$v(z, Z_D) = z \cdot v_z + Z_D \cdot v_D, \tag{12}$$

where $v_z > 0$ and $v_D \geq 0$ are endogenous coefficients, and $v(\cdot)$ solves the following Hamilton-Jacobi-Bellman (HJB) equation,

$$\begin{aligned}(\rho + g)v(z, Z_D) = \max_{\widehat{x},\, \widehat{q},\, \boldsymbol{U}^*} \Big\{ &\pi z + \phi h(z, Z_D) - \delta z\frac{\partial v}{\partial z} + H(z, Z_D)\frac{\partial v}{\partial Z_D} \\ &- wzc_x(\widehat{x}) + z\widehat{x}\Big(- wc_q(\widehat{q}) - \phi c + \mathbb{E}[v(z^{new}(\widehat{q}, \boldsymbol{U}^*), 0) \mid \widehat{q}]\Big)\Big\}.\end{aligned}$$

It is straightforward to verify that coefficients v_z and v_D satisfy the following equations

$$(\rho + g)v_D = \phi h_D + H_D v_D, \tag{13}$$

$$(\rho + g + \delta)v_z = \pi + (\phi h_z + H_z v_D) + \max_{\widehat{x},\, \widehat{q}}\left\{-wc_x(\widehat{x}) + \widehat{x}(-wc_q(\widehat{q}) - \phi c + z^e(\widehat{q})v_z)\right\}. \tag{14}$$

The system of eqs. (13) and (14) suggests that citation-network dynamics affect each firm both as a maker and as a receiver of citations, through the to-be-specified coefficients (h_z, h_D, H_z, H_D) and function $z^e(\cdot)$. The value $\phi > 0$ per citation *received* directly increases the marginal value v_z of patent quality through h_z. If $h_D > 0$, then existing high-quality citations help attract future citations, and a patent's weighted indegree has marginal value $v_D > 0$, which further elevates v_z. High v_z incentivizes firms to pick large $\widehat{q}$ and $\widehat{x}$, *ceteris paribus*. Meanwhile, selection of quality-relevant input patents into $\boldsymbol{U}^*$ determines the return of innovation effort via $z^e(\cdot)$, and formation of $\boldsymbol{U}$ costs an additional ϕc per new patent. Even in the extreme case of $\phi = 0$, innovation decisions still depend on formation of $\boldsymbol{U}^*$ through input quality in $z^e(\cdot)$. Proposition 1 summarizes firm value and decisions given the aggregate state.

Proposition 1 (Firm value and symmetric decisions)**.** *Fix a set of aggregate variables $(g, \overline{z}, w)$, a function $z^e : [\underline{q}, 1] \to \mathbb{R}_+$ for the expected new-patent quality, and coefficients (h_z, h_D, H_z, H_D) for the expected indegree dynamics in eqs. (6) and (7). Eq. (12) is a patent's growth-adjusted value to its owner firm; value coefficients (v_z, v_D) satisfy eqs. (13) and (14). Firms' optimal innovation decisions are patent-wise symmetric, characterized by the common (x, q) that solves eq. (14).*

Larger firms with higher-quality patents innovate more with the symmetric choice of (x, q) *per patent* owned. Given $\boldsymbol{S}$ at t, firm f gets new patents at Poisson rate $x \sum_{\omega=1}^{n} z^{\omega}$, which increases in the number of patents owned and their quality. Achieving such an innovation rate requires the expected flow of research labor hiring to be $\frac{1}{N(t)} \sum_{\omega=1}^{n} z^{\omega} \left(c_x(x) + x c_q(q) \right)$.

Firm value hints at a few externalities with network origins that affect firms' decision margins. For example, input quality in $z^e(\cdot)$ suggests a knowledge externality that interacts with $\boldsymbol{U}^*$'s formation, and the value of a patent owned by one firm contains a $Z_D \cdot v_D$ component contributed by patents owned by other firms.

Equilibrium. We are ready to define the equilibrium formally and list the conditions.

Definition 1. A *stationary equilibrium with balanced growth* consists of a tuple of aggregate variables $(g, \overline{z}, w) \in \mathbb{R}^3_+$, the expected new-patent quality $z^e : [\underline{q}, 1] \to \mathbb{R}_+$ as a function of the chosen step size, coefficients of network dynamics $(h_z, h_D, H_z, H_D) \in \mathbb{R}^3_+ \times \mathbb{R}$, and individual firm decisions summarized by $(x, q) \in [\underline{q}, 1] \times \mathbb{R}_+$ and value function coefficients $(v_z, v_D) \in \mathbb{R}^2_+$, such that they satisfy the following conditions:

1. (v_z, v_D) and (x, q) take into account household optimization, firms' optimal pricing and production, and cleared markets for all products and production labor;

2. given $(g, \overline{z}, w)$, $z^e(\cdot)$, and (h_z, h_D, H_z, H_D), firm optimization yields (x, q) and (v_z, v_D) as summarized by Proposition 1;

3. $(g, \overline{z})$, $z^e(\cdot)$, and firm decisions are consistent, such that the average new-patent quality satisfies $\overline{z} = z^e(q)$, and the aggregate growth rate satisfies $g = x z^e(q) - \delta$;

4. the market for research labor clears such that $R = \langle z \rangle \left(c_x(x) + x c_q(q)\right)$, where $\langle z \rangle$ is in eq. (4), and $R = R^s(w)$ is given in eq. (8);

5. $z^e(\cdot)$, (h_z, h_D, H_z, H_D), and the citation process in Section 5 are consistent, as in Proposition 2 and Proposition 3.

So far, we have taken as given the network dynamics coefficients (h_z, h_D, H_z, H_D) summarizing the expected law of motion of patent indegrees and the expected new patent quality $z^e(\cdot)$ as a function of quality step size. However, as Definition 1 requires, these items are part of the equilibrium, and they must be consistent with firm decisions. To close the model, we need to describe the network formation process in detail to show the last equilibrium condition explicitly.

5 Network Formation and Structure in Equilibrium

This section formalizes the citation-network formation process motivated by the findings in Section 2. A particular feature is that existing citations ease future knowledge flows in similar directions. The process shapes the equilibrium network structure at the patent level and the sector level after aggregation. Details and proofs are in Appendix C.

5.1 Citation-network formation

Firm decisions and a stochastic rule for each new patent to distribute outgoing citation edges jointly drive network formation. One may think of the citation rule as a patent-level knowledge production technology specifying which inputs to use, where to find them, and their shares. Then, innovators decide on the quality step size and select inputs accordingly. The citation rule modifies from a standard statistical network-formation model.[9] In continuous time, we characterize the dynamics of the patent network as the continuous version of a growing network with discrete vertices, in the same spirit of mean-field approximation (henceforth, MFA) or continuum formalism. A difference is that we consider node heterogeneity, i.e., patent quality, similar to Atalay (2013).[10]

At any time t, an existing patent ω in the network has state $(z(\omega,t), \boldsymbol{U}(\omega), \boldsymbol{D}(\omega,t))$. The patent also has a fixed sector classification $j(\omega) \in \{1, \ldots, J\}$. We specify below how every new patent forms the input set $\boldsymbol{U}$ at birth, including the quality-relevant subset $\boldsymbol{U}^*$,

[9] For a basic textbook description of the node copying process, see section 13.5 of Newman (2018). The approach can be traced back to at least Price (1976) and became better-known in economics after Jackson and Rogers (2007).

[10] Atalay (2013) considers fixed fitness level associated with each node and finds exact solutions to degree dynamics and distribution in a discrete network using the master-equation approach. We exploit the model's continuous nature and allow for patent quality depreciation.

conditional on a chosen step size $\widehat{q}$. The functional form of $z^e(\cdot)$ follows. We proceed to derive the law of motion of $\boldsymbol{D}$ captured by functions $h(\cdot)$ and $H(\cdot)$.

Imperfect selection of quality-relevant input patents. Suppose that a new patent is about to select (cite) an existing patent into its quality-relevant input set $\boldsymbol{U}^*$ from a given feasible set $\widetilde{\boldsymbol{N}}$ of patents. Regardless of $\boldsymbol{U}^*$'s size, when choosing each member in it, an innovating firm always has the incentive to choose the highest-quality patent among all feasible choices. However, the choice is imperfect, subject to a random pairwise fitness shifter ι independently drawn from a standard Gumbel distribution with the cumulative distribution function (cdf) given as $\exp(-\exp(-\iota))$. Hence, given a feasible set $\widetilde{\boldsymbol{N}}$ at t, a new patent ω selects $u \in \widetilde{\boldsymbol{N}}$ as an input if it solves $\sup_{\widehat{u}\in\widetilde{\boldsymbol{N}}} z(\widehat{u}, t) \cdot \exp(\iota(\omega, \widehat{u}, t))$, where $\iota(\omega, \widehat{u}, t)$ is the random fitness shifter draw between ω and each $\widehat{u} \in \widetilde{\boldsymbol{N}}$ at t. The citation rule determines the feasible input set $\widetilde{\boldsymbol{N}}$ for each new patent's every quality-relevant citation at t, specified next; it is either the set $\boldsymbol{N}(t)$ of all existing patents or the set $\boldsymbol{N}_j(t)$ of all patents in some specific sector j.

Lemma 1 (Imperfect selection of input quality)**.** *Ex ante, a time-t new patent ω is expected to select each quality-relevant input patent according to a* choice distribution, *such that an existing patent u in the feasible set $\widetilde{\boldsymbol{N}}$ with measure $\widetilde{N} > 0$ gets cited with probability density* $\frac{z(u,t)}{\int_0^{\widetilde{N}} z(\widehat{u},t)\mathrm{d}\widehat{u}}$.

Selection imperfection may also arise from search or informational frictions. We introduce such frictions using very specific functional forms to obtain linear functions $h(\cdot)$ and $H(\cdot)$ to describe the expected indegree dynamics.

U and U*'s formation and the expected new-patent quality. A new patent in sector j at time t is about to cite c existing patents. Its quality step size is $\widehat{q} \in [\underline{q}, 1]$ with $\underline{q} \geq \frac{1}{c}$.

One of the c citations reflects a within-sector parent-child relation. The new sector-j patent (*child*) finds a quality-relevant *parent* within sector j, so the feasible set is $\boldsymbol{N}_j(t)$, and the cited *parent* belongs in $\boldsymbol{U}^*$. The new patent distributes the rest $c-1$ citations by a stochastic rule that depends on $\widehat{q}$. With probability $\frac{c\widehat{q}-1}{c-1}$, the child selects a quality-relevant input patent from $\boldsymbol{N}(t)$, and the cited patent belongs in $\boldsymbol{U}^*$. With the complementary probability $\frac{c(1-\widehat{q})}{c-1}$, the child makes the citation by following its parent's citations. Specifically, the child randomly follows (draws without replacement) one of the parent's citations to a patent u' in sector $j' = j(u')$. Then, with probability $1-\eta \in [0, 1]$, the child copies the citation and cites the same u' regardless of its quality. Since the parent already cites and contains the knowledge of u', this copied citation does not belong to $\boldsymbol{U}^*$. With probability η, the copy *mutates* and the child chooses which one to cite within the followed sector j'. The mutated citation belongs in $\boldsymbol{U}^*$, picked from the feasible set $\boldsymbol{N}_{j'}(t)$.

Figure 2 illustrates the four ways that a new patent distributes citations. Column (a)

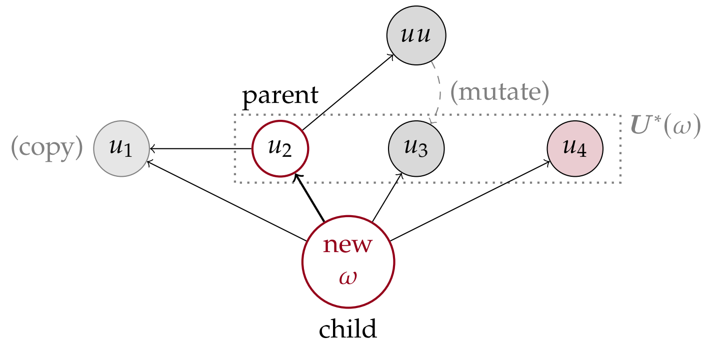


Figure 2: An illustration of the four ways for a newborn node ω to form new edges. New node ω finds and cites a parent ($\omega \to u_2$) within sector. It can independently cite other existing nodes ($\omega \to u_4$), and it may follow its parent's citations ($u_2 \to u_1, uu$) and copy them correctly ($\omega \to u_1$) or with mutations ($\omega \to u_3$).

Note. Nodes plotted in the same style represent patents in the same sector. Citation edges are drawn as solid straight arrows. The thickened edge represents the parent-child connection.

Table 5: Summary of the citation process

	Expected number of citations edges...	
	(a) a *new* patent with $\widehat{q}$ makes	(b) an *existing* patent with (z, Z_D) receives in dt
parent-child $\in \boldsymbol{U}^*$	1	$zg/\langle z\rangle \, \mathrm{d}t$
independent $\in \boldsymbol{U}^*$	$c\widehat{q} - 1$	$z(cq-1)g/\langle z\rangle \, \mathrm{d}t$
exact copies $\in \boldsymbol{U} \setminus \boldsymbol{U}^*$	$c(1-\widehat{q})(1-\eta)$	$Z_D(1-q)(1-\eta)g/\langle z\rangle \, \mathrm{d}t$
mutated copies $\in \boldsymbol{U}^*$	$c(1-\widehat{q})\eta$	$zc(1-q)\eta g/\langle z\rangle \, \mathrm{d}t$
total	c	$[zc(q(1-\eta)+\eta) + Z_D(1-q)(1-\eta)]g/\langle z\rangle \, \mathrm{d}t$

Column (a) shows the expected composition of the input set $\boldsymbol{U}$ of a new patent if it has quality step size $\widehat{q}$ and forms citation edges in four ways according to the citation process described. As a result, Column (b) shows how an existing patent attracts citations in four ways over time when all new patents follow the citation process.

of Table 5 summarizes the expected number of citation edges formed in each way. The quality-relevant input set $\boldsymbol{U}^*$ contains patents selected to form parent-child connection, independent citations, and mutated copies. Each member is chosen from either the set of all patents $\mathcal{N}(t)$ or all patents in some sector $\mathcal{N}_{j'}(t)$, with a choice density specified in Lemma 1. Correct copies of the parent's citations comprise the rest $\boldsymbol{U} \setminus \boldsymbol{U}^*$.

The citation process has plausible features and is tractable, though it may appear complicated. Parent-child connections reflect that innovation in a given research field tends to rely more heavily on existing knowledge in the same field and that a new product likely uses an existing product in the same sector as the base or the blueprint. Independent citations ultimately determine cross-sector citation flows. A new patent making a larger quality improvement is more likely to cite a diverse set of patents in other sectors. Quality step size $\widehat{q}$ thus corresponds to the empirical "*originality*" measure of patent quality. Exact copies of the parent's citations generate the citations-beget-citations dynamics and thus given firms additional incentives to compete for citations. Mutations at rate $\eta \in [0,1]$ capture within-sector citation spillovers in a reduced form. A widely cited patent in a sector brings citations to other patents in the same sector through mutated replications. Crucially, mutation (η) after following a parent's citation ($1-\widehat{q}$) to a sector captures the idea that previous citations pave the ways for subsequent knowledge flows by reducing the cost of new-patent quality. As such, a new patent (*child*) can expand $\boldsymbol{U}^*$ without the cost of $\widehat{q}$ by "free-riding" on its *parent* and following the established citation edges.

The citation process affects the innovation incentives of each firm as a *citer* through the expected new-patent quality $z^e(\widehat{q}) = \mathbb{E}[z^{new}(\widehat{q}, \boldsymbol{U}^*) \mid \widehat{q}]$, which can be written as $\mathbb{E}[z^{new}(\widehat{q}, \boldsymbol{U}^*) \mid \widehat{q}] = \frac{1}{c}\,\mathbb{E}[|\boldsymbol{U}^*| \mid \widehat{q}] \cdot \mathbb{E}[z(u,t) \mid u \in \boldsymbol{U}^*] + \widehat{q}$, where $\mathbb{E}[|\boldsymbol{U}^*| \mid \widehat{q}]$ is the expected size of the quality-relevant input set $\boldsymbol{U}^*$ conditional on a chosen quality step size $\widehat{q}$, and $\mathbb{E}[z(u,t) \mid u \in \boldsymbol{U}^*]$ denotes the expected quality of each input patent selected into $\boldsymbol{U}^*$.

Proposition 2 (Expected new-patent quality)**.** *The citation process is as described, with aggregate variables g and $\overline{z}$. An individual firm expects its new patent's quality to be a function of the quality step size choice $\widehat{q} \in [\underline{q}, 1]$, given as*

$$z^e(\widehat{q};\ g, \overline{z}) = (\widehat{q} + (1-\widehat{q})\eta)\frac{\langle z^2 \rangle}{\langle z \rangle} + \widehat{q}, \tag{15}$$

in which $\mathbb{E}[|\boldsymbol{U}^*| \mid \widehat{q}] = c\widehat{q} + c(1-\widehat{q})\eta$, and $\mathbb{E}[z(u,t) \mid u \in \boldsymbol{U}^*] = \frac{\langle z^2 \rangle}{\langle z \rangle} = \frac{g+\delta}{g+2\delta}\overline{z}$ by MFA.

Two things in Proposition 2 need further discussions. First, given a chosen $\widehat{q}$, the expected size of $\boldsymbol{U}^*$ increases in the within-sector spillover captured by the mutation rate η. In other words, *ceteris paribus*, it is less costly (lower $\widehat{q}$) to achieve a given new-patent quality level when free-riding on the parent is more rewarding. Second, we approximate the expected quality of any selected input patent without considering any within-age-cohort quality differentiation among patents; it is exact if all new patents share the same deterministic quality $\overline{z}$. It is higher than the cross-sectional average quality $\langle z \rangle$ but lower than the mean quality of new patents $\overline{z}$ as selection is imperfect. In the aggregate, given g, the consistency condition $z^e(q) = \overline{z}$ in Definition 1 implies that $\overline{z}$ and q serve as equivalent aggregate variables, such that

$$\overline{z} = q \cdot m(g,q), \quad \text{where } m(g,q) \equiv \left(1 - \frac{g+\delta}{g+2\delta}(q + (1-q)\eta)\right)^{-1}. \tag{16}$$

It is intuitive that the expected new-patent quality $\overline{z}$ is the constant quality step size q amplified by an endogenous multiplier $m(g,q) > 1$ determined by the degree to which existing knowledge helps build new knowledge.

Evolution of D. We turn to the receiving end of citation edges and describe the expected evolution of any given patent's citer set $\boldsymbol{D}(\omega, t)$. Existing patents gain citations as new patents enter. Therefore, evolution of any existing patent ω's $\boldsymbol{D}(\omega, t)$ must be consistent with each newborn patent's citation process.

From t to $t + \mathrm{d}t$ with $\mathrm{d}t$ small, the measure of newborn nodes is approximately $gN(t)\mathrm{d}t$, and they form $gcN(t)\mathrm{d}t$ citation edges with $N(t)$ existing patents. Suppose that these newborn nodes share a common quality step size $q \in [\underline{q}, 1]$. Evolution of $\boldsymbol{D}(\omega, t)$ depends on the patent's own quality $z(\omega, t)$. Furthermore, each existing citer in $\boldsymbol{D}(\omega, t)$ may become a parent of new patents, bringing new citations indirectly via exact copies. Therefore, evolution of $\boldsymbol{D}(\omega, t)$ also depends on the quality-weighted indegree $Z_D(\omega, t)$. Column (b) of Table 5 lists how many new citations an existing patent expects to gain over $\mathrm{d}t$, summarized as follows.

Proposition 3 (Expected law of motion of indegrees)**.** *The citation process is as described, with aggregate variables g, q, and $\overline{z}$. At time t, an existing patent ω has quality z and quality-weighted indegree Z_D. In expectations, ω gains new citations at rate $h(z, Z_D) = \lim_{\mathrm{d}t \to 0^+} \frac{1}{\mathrm{d}t} \mathbb{E}[|\boldsymbol{D}(\omega, t + \mathrm{d}t)| - |\boldsymbol{D}(\omega, t)| \mid z, Z_D] = z \cdot h_z + Z_D \cdot h_D$, where*

$$h_z = \frac{1}{\langle z \rangle} cg(q(1-\eta)+\eta) > 0, \qquad \text{and} \qquad h_D = \frac{1}{\langle z \rangle} g(1-q)(1-\eta) \geq 0. \tag{17}$$

The quality-weighted indegree is expected to evolve as $H(z, Z_D) = \lim_{\mathrm{d}t \to 0^+} \frac{1}{\mathrm{d}t} \mathbb{E}[Z_D(\omega, t + \mathrm{d}t) - Z_D(\omega, t) \mid z, Z_D] = z \cdot H_z + Z_D \cdot H_D$, where

$$H_z = c(q(1-\eta)+\eta)(g+\delta) > 0, \qquad \text{and} \qquad H_D = (1-q)(1-\eta)(g+\delta) - \delta \geq -\delta. \tag{18}$$

Proposition 3 does not require the aggregate g, q, and $\overline{z}$ to satisfy the equilibrium conditions to produce the two functions $h(\cdot)$ and $H(\cdot)$. When these aggregate variables are consistent with one another as in eq. (16), coefficients (h_z, h_D, H_z, H_D) in eqs. (17) and (18) can be viewed as functions of g and q only along an equilibrium path. In continuous time and with an expanding continuum of nodes, $h(\cdot)$ and $H(\cdot)$ exactly describe the expected indegree dynamics. The term $z \cdot h_z$ is the rate at which patent ω attracts new citations by being selected as newborn patents' quality-relevant input, and the term $Z_D \cdot h_D$ is due to new patents' correct copies when members of $\boldsymbol{D}(\omega, t)$ become parents; $H(z, Z_D) = \overline{z}h(z, Z_D) - \delta Z_D$ as every new citation gained is made by a new patent with expected quality $\overline{z}$, and all patents' quality depreciates at δ.

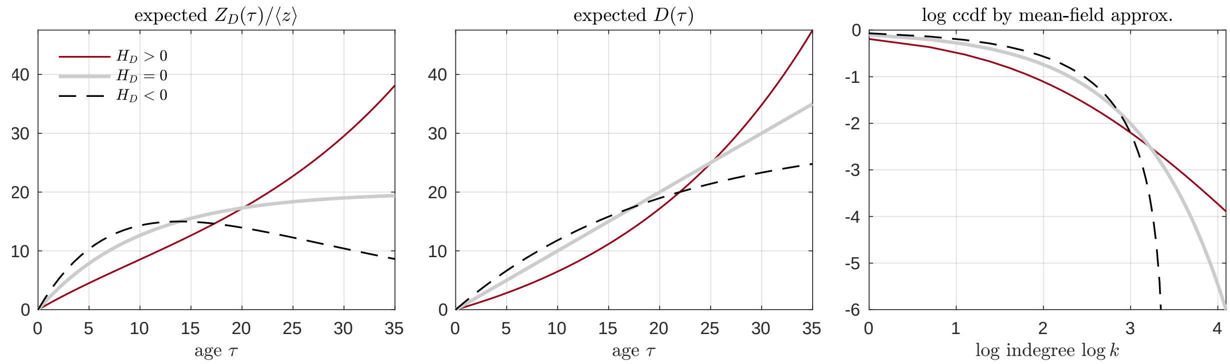


Figure 3: An illustration of the expected indegree dynamics in Corollary 1 and the stationary cross-sectional indegree distribution by MFA in Proposition 4.

Note. The left and middle panels illustrate Corollary 1 at $\widehat{z} = \overline{z}$. The left on plots the expected quality-weighted indegree adjusted by the cross-sectional average quality $\langle z \rangle$, and the middle one plots the expected unweighted indegree. The right panel corresponds to the second item in Proposition 4 and plots the approximated indegree distribution's tail, i.e., the ccdf in log-log scale. The three curves in each panel illustrate three cases of H_D, such that $c = 10$, $g = 0.1$, and $\delta = 0.1$ remain unchanged in each case, whereas the rate of exact copying $(1-q)(1-\eta)$ takes values 0.75, 0.5, and 0.25, respectively.

Corollary 1 (Expected indegree by initial quality and age)**.** *The citation process is as described, with aggregate variables g, q, and $\overline{z}$. A patent has initial quality $\widehat{z}$ at birth. Denote its expected quality-weighted and unweighted indegrees at age $\tau > 0$ as $Z^e_D(\tau;\widehat{z})$ and $D^e(\tau;\widehat{z})$, respectively. At generic H_z and H_D in eq. (18), $Z^e_D(\tau;\widehat{z}) = \widehat{z} H_z \frac{\exp(H_D\tau)-\exp(-\delta\tau)}{H_D+\delta}$, and $D^e(\tau;\widehat{z}) = \widehat{z} h_z \frac{\exp(H_D\tau)-1}{H_D}$. Otherwise, $Z^e_D(\tau;\widehat{z}) = \widehat{z} H_z \tau \exp(-\delta\tau)$ at $H_D = -\delta$, and $D^e(\tau;\widehat{z}) = \frac{\widehat{z}}{\overline{\overline{z}}} c g \tau$ at $H_D = 0$.*

Corollary 1 immediately follows Proposition 3 by solving the differential equations. It establishes the expected relationship between a patent's indegree at a certain age and its initial quality. See the first two panels of Figure 3 for an illustration.

5.2 Equilibrium network structure: patent level

By construction, the network is directed and *acyclic*, with newer patents citing older ones; the network is not locally tree-like due to citation replications. We use the same *directed clustering coefficient* (dcc) as in Section 2 to measure the the stationary network's transitivity. It is the fraction of transitive triplets among all directed paths of length two, i.e., $dcc = \Pr(d \to u \mid d \to \omega \to u)$, where d, ω, u represent patents in the network and "$\to$" a directed citation edge. In this model, dcc has a flavor of redundancy because whenever $d \to \omega \to u$, the additional edge $d \to u$ does not contribute to d's quality.

Also of interest is the shape of the stationary cross-sectional distribution of patent-level citations (indegrees), especially towards the right tail where patents have many citations. We show the counter cumulative distribution function (ccdf) of the stationary patent-level cross-sectional indegree distribution, denoted as $1 - F(k) = \Pr(|\boldsymbol{D}| > k)$. We rely on mean-field approximation to characterize the indegree distribution.

Proposition 4 (Stationary patent-level network structure)**.** *Consider the citation process with growth rate g, quality step size q, and new-patent quality $\overline{z}$. The resulting network at its stationarity has the following features.*

1. *(Transitivity.) The directed clustering coefficient is $dcc = \frac{1}{c}(1-q)(1-\eta) \geq 0$.*
2. *(Indegree distribution.) Let H_z and H_D be the coefficients given in eq. (18), describing the expected law of motion of quality-weighted indegrees. The unweighted indegree distribution's shape depends on the sign of H_D as follows:*
   (a) *if $H_D > 0$, the right tail approaches the power law with $1 - F(k) = \left(\frac{k}{H_z/H_D} + 1\right)^{-g/H_D}$, where the tail exponent $g/H_D > 1$, consistent with a finite mean degree c;*
   (b) *if $H_D = 0$, the right tail is exponential with $1 - F(k) = \exp\left(-\frac{k}{c}\right)$;*
   (c) *if $H_D < 0$, the support is expected to have a finite upper bound $\frac{H_z}{-H_D}$.*

The transitivity measure dcc decreases in quality step size q. Intuitively, lower q means that innovation is more incremental than radical, and a new patent and its parent node share more common citations and thus have more similar knowledge inputs. In other words, some knowledge likely gets repeatedly used as inputs. The model-implied dcc likely *understates* its data counterpart. With a continuum of patent nodes in the model, the only way to form transitive triplets with a positive probability is by correctly copying the parents' citations, but it is not the case in the data.

The stationary indegree distribution's shape depends on how the market allocates innovation resources between quantity (g) and quality (q) purposes in equilibrium. See the right panel of Figure 3 for an illustration. The critical coefficient H_D increases in g and decreases in q; the distribution's shape shifts when H_D switches signs. Quality-weighted indegree Z_D grows at an exponential rate when $H_D > 0$, which ultimately generates the "citations-beget-citations" dynamics. If g is relatively high and q is relatively low, H_D is more likely to be positive, in which case the citation distribution has a power-law right tail, such that a small fraction of "star" patents gets the majority of citations. This is similar to a standard preferential attachment model. A smaller tail exponent g/H_D means a thicker right tail, and it happens when g is higher and q is lower. In the special case of $H_D = 0$, depreciation of existing citer patents' quality cancels out the citations-beget-citations effect. The citation distribution becomes exponential with an ultra-thin right tail, resembling the outcome of a purely random rule to make citations. However, a crucial difference is that the citations-beget-citations effect never disappears when correct copies exist with $\eta < 1$ and $q < 1$, even if $H_D \leq 0$. A negative H_D means that quality of patents in $\boldsymbol{D}$ depreciates too fast relative to their ability to bring more citations in the future. Existing knowledge quickly becomes obsolete, so no patent is expected to attract many citations. The citation distribution's right tail disappears.

5.3 Equilibrium network structure: sector-level aggregation

The patent-level citation network has unweighted edges. Aggregation to the sector level yields a *weighted* digraph, with looping edges representing citations between patents that belong to the same sector. While the patent-level network has increasingly many nodes and edges, the aggregated network has a fixed number of vertices J.

Proposition 5 (Sector-level citation network)**.** *Consider an equilibrium path with quality step size q and common growth rates of all sectors. Let $\mathbf{s} = (s_j)$ be a $J \times 1$ vector summarizing the time-invariant sectoral patent stock shares, such that $s_j = N_j(t)/N(t)$. A symmetric $J \times J$ weighted adjacency matrix $\boldsymbol{\Omega} = [\Omega_{ij}]$ represents the stationary sector-level citation network, given as*

$$\boldsymbol{\Omega} = \boldsymbol{\Omega}(q; \mathbf{s}) \equiv \frac{1}{cq}\mathrm{diag}(\mathbf{s}) + \frac{cq-1}{cq}\mathbf{s}\mathbf{s}^\top, \tag{19}$$

where $\mathrm{diag}(\cdot)$ *creates a diagonal matrix using a given vector, and $\sum_{i,j} \Omega_{ij} = 1$. A typical element Ω_{ij} is the fraction of existing citations from sector i to sector j, and it is also the likelihood that an arbitrary new citation edge goes from i to j.*

Proposition 5 describes cross-sectoral knowledge flows resulting from the patent-level citation process. Symmetry of $\boldsymbol{\Omega}$ naturally follows the absence of "real" heterogeneity among sectors. The matrix $\boldsymbol{\Omega}$ offers several insights regarding sector-level innovations; the implications shed light on more general cases despite its simplicity. First, we observe how sector-level knowledge stock shares enter the weighted adjacency matrix for citation flows. The model does not attempt to explain the heterogeneity in sectoral knowledge stocks and treats $\boldsymbol{s}$ as an exogenous state. Unsurprisingly, sectors with large knowledge stocks tend to attract and make more citations. The functional form of Ω_{ij} resembles gravity between two sectors, proportional to the product of patent numbers in both sectors. Relatedly, Ω_{ij} captures the relative knowledge diffusion rate of new knowledge from sector i to sector j. Knowledge diffusion between larger sectors or within a large sector is generally faster.

Next, we examine and interpret the endogenous weights $\frac{1}{cq}$ and $\frac{cq-1}{cq}$. The additional terms along the diagonal of $\boldsymbol{\Omega}$ reflect a "home bias" that new patents tend to cite more existing ones in the same sector. Specifically, of all citations originating from any sector j, a fraction $\frac{1}{cq} + \frac{cq-1}{cq}s_j$ goes to sector-j patents, and the rest distributes among the other

$J-1$ sectors. In this model, it stems from the child-parent connections that account for $\frac{1}{c}$ of all citations, which gets amplified because new patents make citations by following their parents. Across sectors, the model implies that knowledge diffusion results from innovating firms' decisions to set a quality step size q and actively absorb knowledge from other sectors. Cross-sector knowledge flows increase in the step size q. A subtler point is that these cross-sector flows captured by the endogenous q depend on the within-sector citation spillovers reflected in the exogenous mutation rate η. We discuss it further in the comparative statics analysis in section 6.2.

Corollary 2 (Sector aggregation)**.** *An arbitrarily coarser sectoral classification rule can be written as a $J \times J^a$ aggregation matrix $\boldsymbol{A} \in \{0,1\}^{J^a \times J}$, $J^a < J \in \mathbb{N}$, such that $\boldsymbol{A}^\top \mathbf{1}_{J^a \times 1} = \mathbf{1}_{J \times 1}$. Given $\boldsymbol{A}$, the symmetric $J^a \times J^a$ weighted adjacency matrix after aggregation satisfies $\boldsymbol{\Omega}^a(q; \boldsymbol{A}\mathbf{s}) = \boldsymbol{A}\boldsymbol{\Omega}\boldsymbol{A}^\top$.*

Corollary 2 concerns the extent to which the sector-level citation network's structure depends on aggregation level. It immediately follows Proposition 5 and says that matrix $\boldsymbol{\Omega}$'s form is immune to aggregation. In contrast, disaggregation requires one to consider citation mutations, i.e., exogenous knowledge spillovers, across narrowly defined sectors, so $\boldsymbol{\Omega}$ must adjust accordingly. Nonetheless, a primary takeaway remains that cross-sector knowledge flows are intimately related to innovators' effort captured by q.

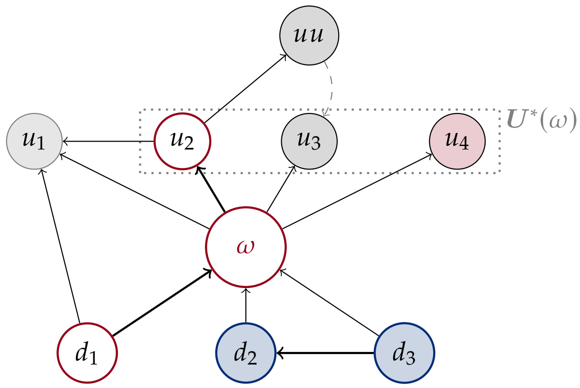
6 Citation Dynamics in Equilibrium and Firm Decisions

The citation process and firm decisions interact with each other in equilibrium. We reexamine the innovating firms' problem to look more closely at the impact of citation dynamics on firm incentives, especially the additional externalities which clearly depend on two critical parameters — knowledge spillover due to mutation η and the value per citation ϕ. We conduct comparative statics analysis to explore the roles of the two parameters.

6.1 Revisiting firm value and decisions: externalities in the network

Armed with an explicit function for expected new-patent quality in Proposition 2 and the coefficients describing indegree dynamics in Proposition 3, we revisit each firm's value and decisions summarized by Proposition 1. The citation process leads to a few kinds of innovation externalities among firms with network origins. Figure 4 serves as an illustrative example for this section.

On the outgoing end of citation edges, the expected new-patent quality $z^e(\widehat{q})$ in eq. (15) as a function of a chosen quality step size $\widehat{q}$ shows jointly a *path externality* and a *level externality* of knowledge. The *level externality* has a familiar aspect that new patents build on the quality level of selected inputs, and input patents have higher expected quality



externality	responsible edges
knowledge: free-ride the past	
level	all $\omega \to \boldsymbol{U}^*(\omega)$ by quality selection
path ($\eta > 0$)	$\omega \to u_3$, a mutated copy of $u_2 \to uu$
value ($\phi > 0$): free-ride the future	
congestion	esp. $d_3 \to \omega$, an exact copy of $d_2 \to \omega$
abundance	all $\boldsymbol{D}(\omega, t) \to \omega$
value sharing	$d_1 \to u_1$, an exact copy of $d_1 \to u_1$

Figure 4: An illustration of the network origins of innovation externalities from patent ω and its owner firm's viewpoint by expanding Figure 1 and Figure 2.

Note. Nodes plotted in the same style represent patents in the same sector. Citation edges are drawn as solid straight arrows. Thickened edges represent parent-child connections.

than the market average due to selection. The novelty is that the degree to which existing knowledge may be exploited is a decision ($\widehat{q}$). Moreover, the decision is under the influence of the *path externality* captured by η. A particular consequence is $\frac{\partial z^e}{\partial \widehat{q}} = 1 + (1-\eta)\frac{g+\delta}{g+2\delta}\bar{z}$. The path externality emerges because new patents follow a $(1-\widehat{q})$ fraction of existing citation edges previously established by their parents and "free ride" on them to select an input at no cost with probability η. In Figure 4, ω "free rides" on u_2 by forming $\omega \to u_3$ as a mutated copy of $u_2 \to uu$. Both kinds of knowledge externalities can be thought of as "*free-riding the past*;" they occur when new patents exploit older ones.

On the receiving end of citation edges, citation dynamics generate a congestion-like effect when the citation value is positive. Network *congestion* results from patents competing for citations in the network with their quality levels. When a firm decides on a quality step size $\widehat{q}$, it does not consider that a larger $\widehat{q}$ elevates the market quality level and makes it more difficult for all patents to attract new citations and to make profits on the product markets. Unlike the competition for profits, having a lead in the competition for citations has an "excess return" due to citations-beget-citations dynamics, which can be viewed as a mirror image of the *path externality*. In Figure 4, every citation such as $d_2 \to \omega$ may subsequently bring $d_3 \to \omega$. More rigorously, observe that the second term of v_z's right-hand side in eq. (14) is due to citation accumulation and can be written as $\phi h_z + H_z v_D = h_z(\phi + \bar{z}v_D) = \phi h_z \frac{\rho+g+\delta}{\rho+g-H_D}$. The first equality says that any increase in a patent's own quality accelerates its citation accumulation by h_z, and every new citation adds flow value ϕ directly and begets a stream of future citations worth $\bar{z}v_D$ as the "excess return" to quality. The second equality, similar to v_D's expression in eq. (13), further illustrates that the weighted indegree Z_D's growth H_D governs the magnitude of said excess return and effectively reduces the discount rate for ϕh_z, with $\rho + g - H_D \leq \rho + g + \delta$. Note that $\frac{\partial h_z(g,q)}{\partial q} < 0$ and $\frac{\partial H_D(g,q)}{\partial q} < 0$. Intuitively, the excess return incentivizes each firm to choose a large quality step size for their patents to be attractive, but it may result in congestion in equilibrium that no patents can be considerably more attractive than the market average; a large q also prevents new patents from copying many of their parents' edges, lowering the excess return.

Citation dynamics also mean a *abundance externality*. Every firm competes for citations from others, and, simultaneously, it innovates and provides c citations to others when a new patent arrives. An individual firm sets the arrival rate of new patents by choosing the innovation intensity $\widehat{x}$. If firms innovates faster and distributes more citation edges, one consequence is that the competition for citations among existing patents is less fierce, reducing network congestion, which is the intuition behind $\frac{\partial h_z(g,q)}{\partial g} > 0$ and $\frac{\partial H_D(g,q)}{\partial g} > 0$. Individual firms do not internalize these consequences when making decisions on $\widehat{x}$.

Last but by no means the least, as both a maker and a receiver of citation edges, every patent has a *value sharing externality*. Specifically, firms have innovation incentives as would-be *owners* of new patents; a new patent with quality $z^e(\widehat{q})$ has an ownership worth of $v(z^e(\widehat{q}), 0) = z^e(\widehat{q})v_z$ in equilibrium. However, a new patent with quality $z^e(\widehat{q})$ cites c existing ones and forms a set $\boldsymbol{U}$, which means that, for each $u \in \boldsymbol{U}$, the new patent becomes a member of u's downstream set $\boldsymbol{D}(u, t)$ and adds $z^e(\widehat{q})v_D$ to u's owner firm's value. In Figure 4, ω is valuable to $\boldsymbol{U}(\omega) = \{u_1, u_2, u_3, u_4\}$'s owners as ω may bring citations such as $d_1 \to u_1$ to them. Therefore, a new patent has a total worth of $v(z^e(\widehat{q}), cz^e(\widehat{q}))$, and its owner cannot claim and thus fails to internalize $cz^e(\widehat{q})v_D$ of it when choosing $\widehat{q}$. In other words, in terms of firm value, every innovating firm enjoys subsequent innovators' effort to improve quality when receiving a citation, and shares its value with predecessors when making one. Indeed, all these value externalities have the flavor of "*free-riding the future*" because existing patents benefit from future patents citing them.

It is worth mentioning that externalities in the network discussed here are not the same as the "*network externality*" in the literature. The latter emphasizes a scale effect such that an action's payoff to an agent increases in the measure of other agents doing the same; the abundance externality discussed above has a similar flavor. Here, the focus is more on the connections in a network than the scale.

Two crucial parameters that jointly determine these externalities are the mutation rate $\eta \in [0, 1]$ and the citation value $\phi \geq 0$. At $\phi = 0$, value externalities are absent; only knowledge externalities remain, and they have the minimal effect at $\eta = 0$ due to lack of any path externality and have the maximal impact at $\eta = 1$ because of the maximal path externality. With $\phi > 0$, the abundance externality appears. Furthermore, value sharing externality and congestion due to citations' excess return reach the maximum when $\eta = 0$. In contrast, no value sharing exists at $\eta = 1$, citations bring proportional return to quality ϕh_z only. The next task is to explore the combined impact of these externalities on the equilibrium outcome in a comparative statics analysis.

6.2 Implications: comparative statics with respect to η and ϕ

This section examines how the equilibrium outcome including the network structure responds to the mutation rate η and the citation value ϕ. The mutation rate η reflects a "constraint" on how existing knowledge is used to produce new knowledge, whereas the citation value ϕ may be subject to policies. Endogenous variables are not always monotone in these parameters, rendering analytical characterizations uninformative. We resort to numerical illustrations instead.

Figure 5 illustrates the model's responses when the within-sector mutation rate η and the citation value ϕ vary. Qualitatively, shade changing when moving up or down shows the effect of η, and shade changing when moving left or right shows the effect of ϕ.

Patent-level citation network. The citation mutation rate η directly governs the production function of new knowledge. Therefore, η's direct impacts on network variables such as h_z and H_D tend to dominate any equilibrium feedbacks, especially when ϕ is small. As η increases over $[0, 1]$, quality-directed selection becomes the primary force of citation dynamics (higher h_z), and the patent-level indegree distribution becomes more concentrated and less transitive (lower dcc). In particular, given a small ϕ, the indegree distribution's right tail becomes thinner (larger g/H_D) and eventually disappears when H_D turns negative; the decline in H_D is due to lowering dcc mainly driven by η. Intuitively, greater path externality reduces citations-beget-citations, and accumulating many citations becomes harder despite the higher entry rate of new patents.

With a given η, the citation value ϕ has general-equilibrium effects on citation dynamics. As ϕ increases with η relatively small, the indegree distribution's right tail also diminishes. However, the decline in H_D in this case is jointly driven by the increase in quality step size q and slowing growth g. Accumulating many citations is harder because new patents enter at a lower rate and make more quality-directed citations.

Sector-level citation network. A particular implication regards how *cross-sector* or global knowledge flows respond to *within-sector* or local spillovers in equilibrium. As η increases, faster growth g means more citations originating from and ending in every sector. However, according to Proposition 5, the coefficient $\frac{cq-1}{cq}$ captures relative knowledge flows across sectors and increases in q, and q decreases in η. More local spillovers are associated with reduced proportions of cross-sector flows and greater "home

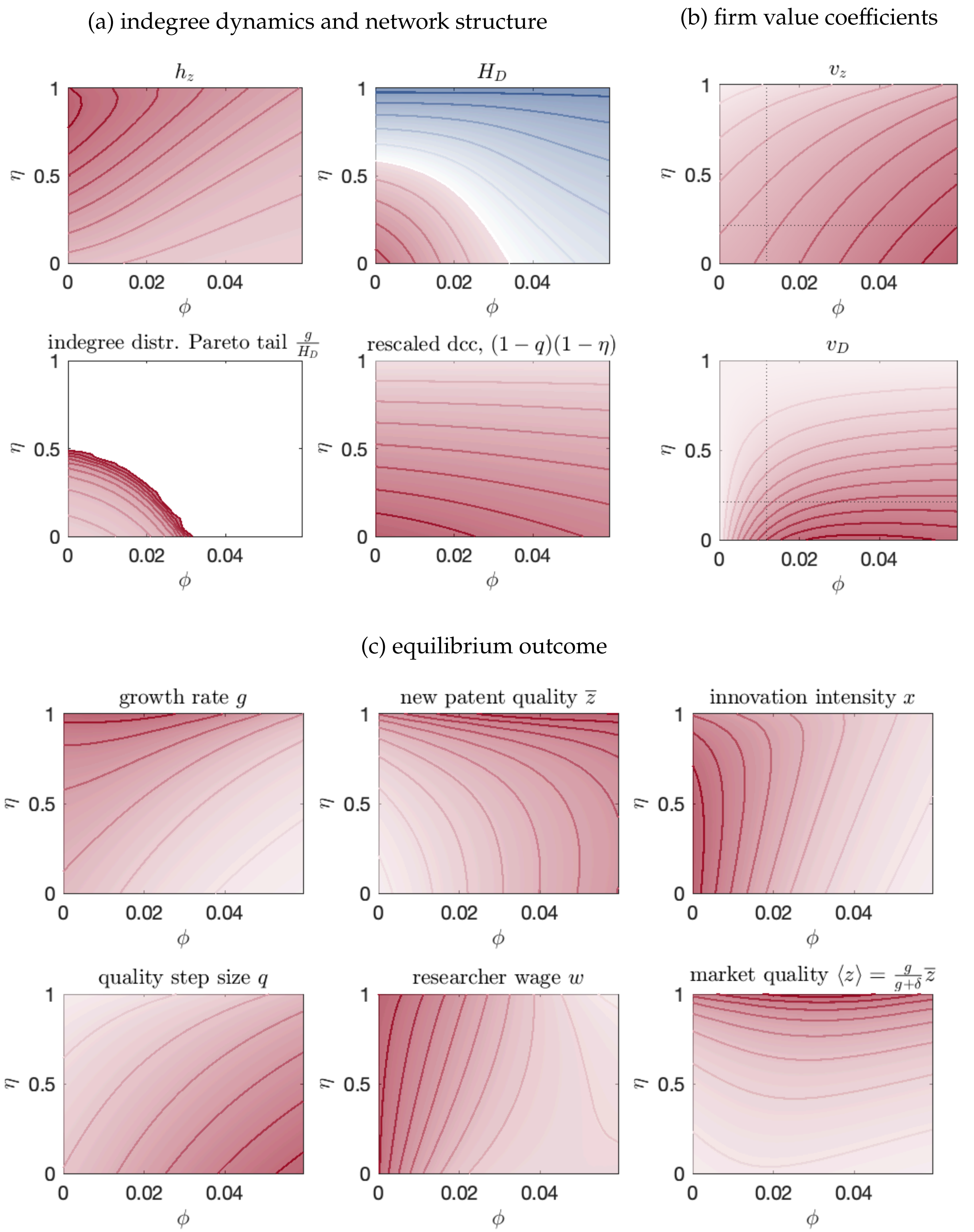


Figure 5: Comparative statics: mutation rate η and citation value ϕ.

Note. In each panel, ϕ is on the x-axis, and η is on the y-axis. Values of each variable of interest are shown in shades (positive in *red* and negative in *blue*) such that *darker* shades represent *higher* absolute values; curves are corresponding contour lines. Blank regions mean undefined values. The two dotted straight lines mark the calibrated values of ϕ and η. All other parameters are at the calibrated values as discussed in Section 7.1.

biases" $\frac{1}{cq}$. That is, on average, although a new patent and its parent are less likely to cite the exact same patents (lower dcc), they are more likely to cite patents in the same sector.

Citation value ϕ may also influence sector-level citations indirectly. Given an η, increases in ϕ drive up q and encourage larger cross-sector knowledge flows relative to within-sector flows, but at the cost of lower growth rate g.

Equilibrium outcome and policy implications. At $\phi = 0$, we have $v_D = 0$, and η affects firm decisions only through knowledge externalities. These externalities remain the dominant channels for changes in η to affect the equilibrium outcome at relatively small $\phi > 0$. As anticipated, when it is easier to follow previously established paths to exploit existing knowledge (higher η), the new-patent quality level $\overline{z}$ is higher, but the size of quality improvement q over inputs is smaller. Meanwhile, growth g gets accelerated if the increase in quality level offsets the possible decline in innovation intensity x, as $g = x\langle z \rangle = x\overline{z} - \delta$. The elevated market quality level $\langle z \rangle$ lowers the profitability coefficient π and the marginal value of quality v_z; it raises the total demand for innovation resources, i.e., research labor R, driving up the equilibrium wage w.

The role of citation value ϕ in determining the equilibrium outcome and its interaction with η shed light on innovation policies. When $\eta \to 1$, the equilibrium outcome becomes less sensitive to ϕ's role as an additional incentive. The reason is that v_D stays close to zero, and increases in ϕ have little effects on v_D due to low excess return of quality; the value sharing externality is also small. Meanwhile, increases in the fixed cost ϕc per new patent discourage innovation intensity x and drag down researcher wage w. Therefore, the observation of q increasing in ϕ is largely driven by the general-equilibrium effect through researcher wage when η is relatively large. When $\eta \to 0$, variations in ϕ have substantial impacts on v_D, so firms' decision margin for q is sensitive to ϕ. As ϕ gets larger, the declining intensity x remains a downward force on wage, but at relatively small η, the downward pressure is partially offset by firms' higher choice of q.

The interaction between ϕ's and η's effects in equilibrium has two policy implications when a government has the tool to influence the effective ϕ. One is the inherent tension between various goals of innovation policies, at a given η. For example, Figure 5c shows that the knowledge-stock growth rate g and the market quality $\langle z \rangle$ achieve their maxima at different ϕ values. An "optimal" choice of ϕ crucially hinges on how the policymakers tradeoff aggregate growth g and average quality $\langle z \rangle$, among other things.[11] Another implication is that the efficacy of these policies depends on the nature of knowledge

[11] If the policymakers aim to maximize household welfare as a benevolent social planner would, then the growth-versus-quality tradeoff depends on the households' degree of love-of-variety, which in turn determines the optimal policy. See Appendix D for details.

Table 6: Parameter Values

predetermined			matched to citation moments			matched to equilibrium moments				
ρ	L	σ	c	δ	η	$\overline{c_x}$	$\overline{c_q}$	ϕ	ν	χ
0.03	1	2	11.48	0.108	0.212	0.603	38.18	0.012	36.98	43.87

production, and so does the optimal policy choice. As is previously discussed, η's value governs whether the equilibrium outcome is sensitive to ϕ. Furthermore, with a given objective of policymakers, the optimal choice of ϕ differs when η changes. We explore this numerically in the upcoming section.

7 Quantitative Explorations

We briefly discuss the model's calibration procedure and numerically demonstrate the policy implications. The model is highly stylized for tractability, and the quantitative exercise is illustrative. The goal is to show that the knowledge and value externalities in the citation network considerably impact the optimal policy and its efficacy.

7.1 Parameter calibration

We calibrate the model along a balanced-growth equilibrium path. Table 6 reports the parameter values. The unit length of time is a year, and the discount rate is $\rho = 0.03$. We normalize the supply of production labor (numeraire) to be $L = 1$. The supply elasticity of research labor is set at 0.5, i.e., $\sigma = 2$. Cost functions are quadratic such that $c_x(x) = \frac{1}{2}\overline{c_x}x^2$ and $c_q(q) = \frac{1}{2}\overline{c_q}(q - \underline{q})^2$ with $\underline{q} = \frac{1}{c}$. Parameters $\overline{c_x}$, $\overline{c_q}$, and c remain to be set.

The primary data sources that we use to pin down the rest of the parameter values are the USPTO citation data and the Business Enterprise Research and Development Survey (BERD) from the National Science Foundation (NSF). The former data set helps to back out network-related variables and parameters. The average citations made per patent c has an immediate empirical counterpart. To set the values of mutation rate η and quality discount rate δ, we exploit the citation network dynamics and structure according to Proposition 3, Proposition 4 and Proposition 5 as follows. The network growth rate g is set at the aggregate birth rate of new patents. Then, we calculate the quality step size q using cross-sector citations flows as shown in Proposition 5. With these variables at hand, we do not need to simulate the model to find η and δ. Citation dynamics described in Proposition 3 rely on η and δ. We aggregate patents by birth-year cohort and look at the total number of new citations attracted by each cohort in each sample year. These cross-cohort and over-time variations become one set of empirical moments to discipline the

Table 7: Goodness of fit

targeted moments	data	model
new patent quality step size, q	0.223	0.223
aggregate innovation rate, g	0.129	0.129
researcher compensation as a fraction of total R&D, $wR/(wR + g\phi c)$	0.526	0.526
total R&D as fraction of net sales, $(wR + g\phi c)/(L\nu/(\nu - 1))$	0.036	0.036
researcher employment share, R/L	0.076	0.076

The sample period is from 1976 to 2014. All data moments are the means of the corresponding time series. The fraction of patenting firm-sectors approximates innovation probability. The new patent quality calculation follows the procedure described in the main text using USPTO citation data. Aggregate innovation rate is the growth rate of patent counts. Researcher compensation fraction and R&D to sales ratio are based on NSF BERD data. Researcher compensation includes wages and those in other forms such as fringe benefits. Other costs, such as royalties and supplies, are considered payments between firms.

two parameters. Proposition 4 implies that the patent-level citation distribution's shape are also informative about η and δ. In the data, this distribution has an approximate power-law right tail, and we estimate the tail exponent year by year following Clauset, Shalizi, and Newman (2009). The estimated tail exponent remains stable over the sample period and is approximately 3.5, which is another moment for η and δ. Thus, we are able to pin down the two parameter values. The two endogenous variables g and q obtained in the interim are also useful in the next step. Details of the procedure is in Appendix E.

We pick the remaining five parameters in the last panel of Table 6 using the simulated method of moments (SMM), such that the model matches a set of five chosen moments. These moments include previously obtained aggregate innovation or growth rate g and equilibrium quality step size of new patents q, as well as the total researcher compensation as a fraction of all R&D expenditures, the R&D expenditure as a fraction of net sales, and the ratio between research labor and production labor R/L. The last three moments are calculated using the NSF BERD data. Table 7 show that the model fully reproduces these moments. Identification intuitions are as follows. The quality step size q is a direct target for $\overline{c_q}$, and the aggregate growth rate becomes a direct target for $\overline{c_x}$. In equilibrium, the total R&D expenditure aggregated across all firms and sectors is $wR + g\phi c$. We interpret wR as all the compensations to researchers in the data and match $g\phi c$ to other types of cost in the data, which typically involve payments among firms. Hence, the researcher compensation fraction is a direct target for ϕ. The substitution elasticity ν governs the aggregate sales revenue, so the R&D to sales ratio is a direct target for ν. Lastly, researcher employment share regulates the level coefficient χ of research labor supply.

7.2 Knowledge externality and the optimal innovation policy

This section continues the discussion in section 6.2 on innovation policies. Consider a government capable of setting the citation value ϕ directly. Depending on the sign of the

Table 8: Optimal $\phi^\star$ depends on citation spillover rate η and household love-of-variety

preference	η		welfare gap ($\Delta C\%$)	growth g	mkt. quality $\langle z \rangle$	researchers R
full love-of-variety	0	first-best		0.230	0.146	0.151
(growth over quality)		market, $\phi^\star = -0.025$	2.71	0.170	0.173	0.097
	0.212	first-best		0.257	0.173	0.157
		market, $\phi^\star = -0.011$	4.09	0.171	0.128	0.086
		baseline market	7.04	0.129	0.165	0.076
	1	first-best		1.045	1.402	0.277
		market, $\phi^\star = 0.010$	38.7	0.332	0.471	0.080
no love-of-variety	0	first-best		0.079	0.146	0.071
(quality over growth)		market, $\phi^\star = 0.021$	0.000	0.079	0.146	0.072
	0.212	first-best		0.090	0.167	0.071
		market, $\phi^\star = 0.025$	0.001	0.087	0.169	0.072
		baseline market	0.155	0.129	0.165	0.076
	1	first-best		0.287	0.562	0.079
		market, $\phi^\star = 0.036$	0.025	0.269	0.524	0.072

The welfare gap is reported in percentage points as the additional equilibrium consumption needed at every t for the households' lifetime utility to achieve the first-best level. The optimal $\phi^\star$ is the value of ϕ that maximizes the households' lifetime utility in equilibrium given other parameters.

chosen ϕ, such a simplified policy tool is equivalent to a tax-and-subsidy policy combination that alters firms' relative decision margins of quality step size $\widehat{q}$ and innovation intensity $\widehat{x}$, and the government budget is always balanced. Suppose that the government aims to maximize household welfare by choosing the optimal $\phi^\star$.

Table 8 shows that the optimal policy $\phi^\star$ and its efficacy depend on local knowledge spillover η and the government's policy objective determined by household preferences. In particular, the degree of households' love-of-variety governs the quality-versus-quantity tradeoff, and the first-best outcome in each scenario is a benevolent social planner's choice. See Appendix D for details.

With love-of-variety, having more varieties in the economy improves household welfare directly. Consequently, the planner's objective and the government's goal have greater emphasis on growth g than the market quality $\langle z \rangle$. Meanwhile, household love-of-variety is yet another source of externality that the production side fails to internalize. Therefore, it is intuitive that altering transfers among firms by choosing $\phi^\star$ cannot fully restore the social optimum. The optimal value and sign of $\phi^\star$ and how well it works to narrow the welfare gap rely on the nature of knowledge production. At the extreme of no path externality in knowledge production, i.e., $\eta = 0$, improving quality level is relatively costly for the planner. The optimal $\phi^\star < 0$ means taxing the quality ($\widehat{q}$) margin while subsidizing the quantity margin ($\widehat{x}$). In this case, value externalities become negative. The remaining welfare gap between the market equilibrium and the social optimum is relatively small. When $\eta > 0$ is low such as the calibrated value, the path externality is small, and the optimal policy and its efficacy are qualitatively the same. As $\eta \to 1$, the optimal policy becomes less effective in narrowing the gap between the market equilibrium and the social optimum. The knowledge externality, household love-of-variety, and the lack of any excess return to quality are all at play.

Without love-of-variety, the additional source of externality is absent, and the optimal policy is much more effective. The planner and the government emphasize the market quality $\langle z \rangle$ rather than growth g. At $\eta = 0$, value externalities are at the maximum but path externality does not exist. The optimal $\phi^\star > 0$ means rewarding the quality margin while taxing the quantity margin, and it almost fully removes the discrepancy between the social optimal and the market equilibrium. Even as $\eta \to 1$, the optimal policy remains effective in narrowing the welfare gap. The reason is that ϕ affects firms' decision margins on quality relative to quantity directly and in the opposite directions. Consequently, it is effective in reallocating research labor between quality and quantity purposes, and less so in driving the overall demand for research labor. Without love-of-variety, the level of research labor at the social optimum and that at the market equilibrium are similar, which is in contrast with the case of full love-of-variety.

Despite the apparent difference between the two cases of household preferences, a common theme remains that the knowledge path externality should be a concern of innovation policies. Another implication is that it is possible to fight (knowledge) externalities with (value) externalities. However, it again requires further understanding of the knowledge production process.

8 Concluding Remarks

We propose and analyze a dynamic model of firm innovation in which citations are path-dependent and generate values for the cited firms. The citation process is motivated by empirical findings and is modeled as a hybrid of endogenous formation and a stochastic process for tractability. The presence of citation value also has support from the data, but the model treats it as exogenous and is silent on its origin. New externalities arise from such a setting, leading to new insights about innovation policies.

As a first attempt to discuss innovation incentives with endogenous citation network dynamics, the framework is intentionally stylized to serve as a baseline. It has a few natural directions of extensions and enrichments for more quantitative-oriented analysis. One is to consider patent-level or firm-level turnovers induced by forces such as creative destructions. These richer dynamics have further implications to discipline the model, such as firm dynamics and distributions. The network formation process needs adjusting to reflect the turnovers. Another is to introduce richer heterogeneity. For example, the strong assumption of symmetric sectors can be dropped to allow for "real" heterogeneity at the sector level, such as household preferences, production technologies, and barriers to new firm entries. Alternatively, patent quality can be multi-dimensional, such as distinguishing each patent's scientific value and market appeal. Theoretically, one can give the model further micro-foundations and investigate the origin of citation values. For example, it is possible to introduce imperfect information such that patent quality is unobservable to non-researcher agents, so citations are noisy signals about patent quality. We leave these to future research.

References

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. 2018. "Innovation, reallocation, and growth." *American Economic Review* 108 (11):3450–91.

Acemoglu, Daron and Pablo D Azar. 2020. "Endogenous production networks." *Econometrica* 88 (1):33–82.

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt. 2015. "The Schumpeterian Growth Paradigm." *Annual Review of Economics* 7 (1):557–575.

Aghion, Philippe and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60 (2):323–351.

Akcigit, Ufuk, Murat Alp Celik, and Jeremy Greenwood. 2016. "Buy, keep, or sell: Economic growth and the market for ideas." *Econometrica* 84 (3):943–984.

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde. 2020. "Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth." *The Review of Economic Studies* 88 (1):1–43.

Akcigit, Ufuk and William R Kerr. 2018. "Growth through heterogeneous innovations." *Journal of Political Economy* 126 (4):1374–1443.

Arora, Ashish, Wesley Cohen, Honggi Lee, and Divya Sebastian. 2023. "Invention value, inventive capability and the large firm advantage." *Research Policy* 52 (1):104650.

Atalay, Enghin. 2013. "Sources of variation in social networks." *Games and Economic Behavior* 79:106–131.

Atalay, Enghin, Ali Hortacsu, James Roberts, and Chad Syverson. 2011. "Network structure of production." *Proceedings of the National Academy of Sciences* 108 (13):5199–5202.

Austin, David H. 1993. "An event-study approach to measuring innovative output: The case of biotechnology." *The American economic review* 83 (2):253–258.

Balsmeier, Benjamin, Mohamad Assaf, Tyler Chesebro, Gabe Fierro, Kevin Johnson, Scott Johnson, Guan-Cheng Li, Sonja Lück, Doug O'Reagan, Bill Yeh et al. 2018. "Machine learning and natural language processing on the patent corpus: Data, tools, and new measures." *Journal of Economics & Management Strategy* 27 (3):535–553.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying technology spillovers and product market rivalry." *Econometrica* 81 (4):1347–1393.

Bryan, Kevin A and Jorge Lemus. 2017. "The direction of innovation." *Journal of Economic Theory* 172:247–272.

Bryan, Kevin A and Heidi L Williams. 2021. "Innovation: market failures and public policies." *Handbook of Industrial Organization* 5 (1):281–388.

Caballero, Ricardo J. and Adam B. Jaffe. 1993. "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth." *NBER Macroeconomics Annual* 8:15–74.

Cai, Jie and Nan Li. 2019. "Growth through inter-sectoral knowledge linkages." *The Review of Economic Studies* 86 (5):1827–1866.

Cai, Jie, Nan Li, and Ana Maria Santacreu. 2022. "Knowledge diffusion, trade, and innovation across countries and sectors." *American Economic Journal: Macroeconomics* 14 (1):104–45.

Carvalho, Vasco M and Nico Voigtländer. 2014. "Input diffusion and the evolution of production networks." Tech. rep., National Bureau of Economic Research.

Chaney, Thomas. 2014. "The network structure of international trade." *American Economic Review* 104 (11):3600–3634.

Clauset, Aaron, Cosma Rohilla Shalizi, and Mark EJ Newman. 2009. "Power-law distributions in empirical data." *SIAM review* 51 (4):661–703.

Demir, Banu, Ana Cecília Fieler, Daniel Xu, and Kelly Kaili Yang. 2021. "O-Ring Production Networks." Tech. rep.

Dhyne, Emmanuel, Ayumu Ken Kikkawa, Magne Mogstad, and Felix Tintelnot. 2020. "Trade and Domestic Production Networks." *The Review of Economic Studies* 88 (2):643–668.

Eeckhout, Jan and Boyan Jovanovic. 2002. "Knowledge spillovers and inequality." *American economic review* 92 (5):1290–1307.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow. 2019. "How Destructive Is Innovation?" *Econometrica* 87 (5):1507–1541.

Ghassibe, Mishel. 2020. "Endogenous Production Networks and Non-Linear Monetary Transmission." In *Technical Report*.

Grassi, Basile. 2018. “IO in IO: Size, Industrial Organization, and the Input-Output NetworkMake a Firm Structurally Important.” Tech. rep., IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

Griliches, Zvi. 1957. “Hybrid corn: An exploration in the economics of technological change.” *Econometrica, Journal of the Econometric Society* :501–522.

Grossman, Gene M. and Elhanan Helpman. 1991. “Quality Ladders in the Theory of Growth.” *The Review of Economic Studies* 58 (1):43–61.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg. 2005. “Market value and patent citations.” *RAND Journal of economics* :16–38.

Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg. 2001. “The NBER patent citation data file: Lessons, insights and methodological tools.” Tech. rep., National Bureau of Economic Research.

Helpman, Elhanan. 1998. *General purpose technologies and economic growth*. MIT press.

Higham, Kyle, Gaétan De Rassenfosse, and Adam B Jaffe. 2021. “Patent quality: Towards a systematic framework for analysis and measurement.” *Research Policy* 50 (4):104215.

Higham, Kyle W, Michele Governale, AB Jaffe, and Ulrich Zülicke. 2019. “Ex-ante measure of patent quality reveals intrinsic fitness for citation-network growth.” *Physical Review E* 99 (6):060301.

Hopenhayn, Hugo and Francesco Squintani. 2021. “On the Direction of Innovation.” *Journal of Political Economy* :000–000.

Huang, Jingong Jim et al. 2018a. “Technology Network, Innovation and Growth.” In *2018 Meeting Papers*, 178. Society for Economic Dynamics.

———. 2018b. “Technology Spillover as a Determinant of Merger and Acquisition Decisions.” In *Working Papers*. University of Nottingham, Ningbo.

İmrohoroğlu, Ayşe and Şelale Tüzel. 2014. “Firm-level productivity, risk, and return.” *Management Science* 60 (8):2073–2090.

Jackson, Matthew O. 2010. *Social and economic networks*. Princeton university press.

Jackson, Matthew O and Brian W Rogers. 2007. “Meeting strangers and friends of friends: How random are social networks?” *American Economic Review* 97 (3):890–915.

Jovanovic, Boyan and Rafael Rob. 1990. “Long Waves and Short Waves: Growth Through Intensive and Extensive Search.” *Econometrica* 58 (6):1391–1409.

Jovanovic, Boyan and Peter L Rousseau. 2005. “General purpose technologies.” In *Handbook of economic growth*, vol. 1. Elsevier, 1181–1224.

Klette, Tor Jakob and Samuel Kortum. 2004. "Innovating firms and aggregate innovation." *Journal of political economy* 112 (5):986–1018.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological innovation, resource allocation, and growth." *The Quarterly Journal of Economics* 132 (2):665–712.

Lentz, Rasmus and Dale T Mortensen. 2008. "An empirical model of growth through product innovation." *Econometrica* 76 (6):1317–1373.

Lim, Kevin. 2017. "Firm-to-firm trade in sticky production networks." In *2017 Meeting Papers*, vol. 280. Society for Economic Dynamics.

Liu, Ernest and Song Ma. 2021. "Innovation Networks and Innovation Policy." Tech. rep., National Bureau of Economic Research.

Marx, Matt and Aaron Fuegi. 2020. "Reliance on science: Worldwide front-page patent citations to scientific articles." *Strategic Management Journal* 41 (9):1572–1594.

Newman, Mark. 2018. *Networks*. Oxford university press.

Nicholas, Tom. 2008. "Does innovation cause stock market runups? Evidence from the great crash." *American Economic Review* 98 (4):1370–96.

Oberfield, Ezra. 2018. "A theory of input–output architecture." *Econometrica* 86 (2):559–589.

Olmstead-Rumsey, Jane. 2020. "Market Concentration and the Productivity Slowdown." Tech. rep., Northwestern University.

Pakes, Ariel and Mark Schankerman. 1984. "The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources." In *R&D, patents, and productivity*. University of Chicago Press, 73–88.

Price, Derek De Solla. 1976. "A general theory of bibliometric and other cumulative advantage processes." *Journal of the American Society for Information Science* 27 (5):292–306.

Romer, Paul M. 1986. "Increasing returns and long-run growth." *Journal of political economy* 94 (5):1002–1037.

———. 1990. "Endogenous technological change." *Journal of political Economy* 98 (5, Part 2):S71–S102.

Taschereau-Dumouchel, Mathieu. 2020. "Cascades and fluctuations in an economy with an endogenous production network." *Available at SSRN 3115854* .

Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe. 1997. "University versus corporate patents: A window on the basicness of invention." *Economics of Innovation and new technology* 5 (1):19–50.