Innovation and Endogenous Knowledge Network Dynamics*

Jie Cai† Can Tian‡

July 2024

Abstract

We study innovation in evolving knowledge networks and its growth implications. Empirically, a new link in the knowledge network is approximated as the first citation between a pair of technology categories. We find that these new links likely connect high-quality innovations and large firms on both ends. After forming a new link, responsible firms' innovation and real performance improve on both ends with positive spillovers to other firms. We build a tractable model to rationalize these findings and analytically characterize the stationary equilibrium. In the model, the knowledge network evolves endogenously as each firm searches for others as knowledge input suppliers and attracts others as users of its own knowledge. In equilibrium, faster growth is associated with denser and better knowledge networks. Endogenous link formation amplifies the effect of innovation incentives on growth. The calibrated model suggests that the majority of growth in the US has network origins.

Keywords: innovation, patent citations, endogenous network dynamics, endogenous growth, search and matching

*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, and Daniel Xu for valuable comments on earlier versions of this paper. We thank Murat Celik, Matt Jackson, Samuel Kortum, Erzo Luttmer, Guillermo Ordoñez, Diego Restuccia, Mark Rempel, and Alireza Tahbaz-Salehi for insightful discussions shaping the current version. We also thank participants of many seminars and conference sessions. Any remaining errors are our own.

†School of Economics, Shanghai University of Finance and Economics, Shanghai, 200433, China. april.cai@gmail.com

‡Department of Economics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. tiancan@email.unc.edu

1 Introduction

Knowledge network is crucial for innovation, in the same way that input-output network determines production. It is already established that the disaggregated knowledge network structure, even if taken as fixed, matters for aggregate economic growth. A natural next question is whether and how the knowledge network changes over time in the data and to study the individual incentives behind the network dynamics and the aggregate implications in a model. However, there has been little documentation of knowledge network evolution, and a tractable model suitable for this purpose does not yet exist. This paper is a first step in these directions. We document a set of such empirical findings and propose such a theoretical framework to show that endogenous knowledge network dynamics drive aggregate growth.

To formalize the dynamics of knowledge networks theoretically, we need to understand innovating firms' incentives to find new applicable knowledge inputs in various technology categories. Since the cost of searching in never-before-used categories must be dramatically higher than searching within categories known to be applicable, the potential benefit must be equally large. Therefore, our starting point is to identify the very firms and patents that build the brand new citation linkages across technology categories using patent citation data, and then compare them with average firms and patents in the USPTO patent data base.

First of all, the creation of new citation linkages across different technology categories is a rare event in the USPTO patent citation data. Less than 1% of patent applications includes a citation to a never cited category after 1976; this possibility reduces to less than 0.1% after 2000. We call the firms and patents that make the first citation from a technology category to another category in USPTO history pathfinder firms and patents, and those firms and patents that receive the first citation from another technology category path-receiver firms and patents. These path builders turn out to be extraordinary firms and patents in many dimensions. The path builder firms are much larger in patent stock and scope, and those public path builder firms' financial performances, such as capital stock, employment, market value, sales, profit and TFP, are much better than peer firms, controlling for firm, industry and year fixed effects. The path builder patents score higher in several quality measures, such

as citations to scientific literature, grant lag, cited patents age, number of technology categories granted, pedigree, originality, generality, and the number of forward citations received, than average patents, controlling for category and application year fixed effects. These facts suggest that firms that are more efficient at innovation and firms that provide higher quality knowledge inputs are more likely to connect to each other, and they jointly create more value when building new knowledge links to compensate for the high cost associated with searching in a new knowledge source. As a result, we observe the better performances of path builder firms and patents.

The next question is how firms gain from knowledge pathway creation. Do pathfinders pocket the gains from new link formation entirely or do they share them with the path-receivers? Do peer firms receive any spillovers from such events? We find that not only pathfinders and path-receivers, but also the peer firms that innovate in the same involved technology categories, experience higher growth rates in patent applications, capital, employment, sales, profit, market value and TFP. Such empirical evidences point out that the event of knowledge network formation reveals the high qualities of both pathfinder and path-receiver patents, and the two path builders share the gains from connecting two previously isolated categories. Pathway citations are publicly observed once patent applications are granted, so the peer firms in both the knowledge using and used sides enjoy positive spillovers by following these pathways when innovating and sharing gains in the future.

Given these facts, we want to assess the role of knowledge network dynamics in driving aggregate economic growth. Doing so requires a theoretical model with heterogeneous firms, their innovation decisions, and endogenous knowledge network dynamics. However, an off-the-shelf model with these features does not yet exist due to obvious technical challenges or potential computational costs. Our technical contribution is thus a tractable and flexible framework built with rather standard ingredients. In particular, we consider network formation and evolution with a *continuum* of nodes, each forming at most a *finite* number of connections with other nodes due to search frictions. It allows for aggregate-level stationarity while preserving idiosyncratic and evolving granularity, characterized in closed form.

The main idea of the model is as follows. Each firm has an piece of knowledge and a fixed two-dimensional efficiency type: an *output* efficiency type governing the firm's

own innovation outcome and an *input* efficiency type determining the contribution to others' innovation outcome when its knowledge is used as an input. Each firm plays two roles in innovation. As a knowledge user, a firm must find knowledge inputs owned by others before using them for subsequent innovations, and search takes time and is costly. Meanwhile, as a knowledge input supplier, the firm attracts other firms as users. A knowledge network is thus formed by firms searching for suppliers and attracting users among themselves. The links are naturally directed, reflecting the (inverse) directions of knowledge flows, and the links are also weighted by pairwise *match qualities* determined by the user's output type and the supplier's input type. A firm's profitability jumps up at the time of a successful innovation, and the innovating firm shares the resulting surplus with all of its input suppliers. The gains from innovation incentivize firms to *search* for input sources and to exert efforts to compete for user attention and *attract* new users, thereby encouraging the formation of new connections. Firms with higher *output* types have greater search incentives, and firms with higher *input* types have greater attraction incentives. As a result, the network always evolves as new links are endogenously formed while existing links get randomly and exogenously destroyed.

Network connections are formed at the industry level, leading to positive spillovers if multiple firms innovate in an industry, referred to as the *knowledge path externality*. When one firm in industry j finds another firm in industry i as an input supplier, a j to i pathway is formed. Then, because firms in one industry own similar knowledge inputs and can use similar inputs, all firms in j have i as an input source, and all firms in i have those in j as potential users. Therefore, firms in the same industry free ride on each other's search and attraction efforts to form connections.

The model is tractable despite its rich heterogeneity and dynamics. It captures the empirical findings and allows for analytical characterizations of the stationary equilibrium path. We highlight two major takeaways out of the equilibrium implications. First, in the aggregate, knowledge network dynamics drive growth. Equilibrium growth is faster when innovating firms form a knowledge network with more links and higher match quality per link. In other words, growth is faster if higher-quality knowledge inputs are used more frequently for innovation and if more capable innovators have access to more of these inputs. Moreover, whenever firms have greater incen-

tives to innovate, they also have greater incentives to form new links in the knowledge network, which is an amplifying effect on growth.

Second, to each individual firm, both the existing network connections and the option to gain more connections in the future have values. Existing links to suppliers are valuable because their knowledge inputs can be repeatedly used to generate innovation surplus. Existing links to users are valuable because they generate streams of future rents extracted from users. Firms are thus willing to search for new suppliers and attract new users at a cost of innovation resources, giving rise to the option value of future connections. It follows that firms' incentives to form network connections have nontrivial impact on the within- and cross-firm resource allocation.

Our model casts knowledge diffusion in a new light. Across industries, firms actively search for suitable knowledge inputs owned by other firms in the presence of search frictions, resulting in gradual knowledge flows much like the flows of input goods. Within an industry, diffusion occurs as firms follow the knowledge paths found by others and use new inputs.

Our model provides a convenient decomposition to examine the origins of aggregate economic growth. We calibrate the model to the US economy and show that the knowledge network across technology categories accounts for more than one half of all economic growth in the US. In turn, the knowledge path externality where firms free-ride on others' efforts to form links accounts for the vast majority of growth with network origins. Meanwhile, changes in growth with network origins explain about two thirds of the growth decline in the last few decades.

Layout. After the upcoming literature review, the remainder of the paper proceeds as follows. Section 2 contains data description and empirical findings. Section 3 describes the benchmark model without knowledge path externality. Section 4 characterizes the stationary equilibrium and discusses the implications. Section 5 extends the benchmark model to introduce the knowledge path externality. Section 6 calibrates and quantitatively illustrates the extended model. Section 7 concludes. All proofs and omitted details of the benchmark model are in Appendix A, and those of the extended model are in Appendix B. Details of identification and quantitative analysis are in Appendix C. Appendix D contains additional empirical results.

Related literature. This paper builds on and contributes to several strands of

literature. Network dynamics aside, the model intends to stand close to the literature of endogenous growth and innovating firms such as Klette and Kortum (2004). See, e.g., Aghion et al. (2015) for a comprehensive review. A few papers in this literature study cross-sector knowledge flows in long-term growth and trade, such as Helpman (1998), Jovanovic and Rousseau (2005) Akcigit et al. (2016), Cai and Li (2019), Cai et al. (2022), Liu and Ma (2024). These studies take the knowledge network as given and fixed, whereas the network's endogenous dynamics are our primary focus.

Our theoretical model contributes to the growing literature on endogenous networks in macroeconomics and trade. Examples are Oberfield (2018), Grassi (2018), Lim (2017), Huneeus (2018), Acemoglu and Azar (2020), Taschereau-Dumouchel (2020), Ghassibe (2020), Dhyne et al. (2020), Acemoglu and Tahbaz-Salehi (2023), Kopytov et al. (2023). In particular, Oberfield (2018) is the first to model an endogenous production network with a *continuum* of nodes and locally discrete links. Our model builds on the search and matching framework, a feature shared by a few recent contributions. Arkolakis et al. (2023) and Demir et al. (2024) study static search models in which one firm's production requires a continuum of other firms as input suppliers. Aekka and Khanna (2024) ventures into a dynamic setting where each newborn firm gradually builds the continuum of suppliers over time, treating network evolution as capital accumulation. Panigrahi (2021), Lenoir et al. (2023), and Eaton et al. (2022) allow for idiosyncratic granularity as the random matching outcome in static settings to study firm-to-firm trade, such that a firm trades with at most a finite number of counterparties; the idea is extended into a dynamic setting by Miyauchi (2023). These contributions study buyer-seller networks absent growth implications.[1] Our model features endogenous and heterogeneous search intensities and free-riding effects in dynamically formed knowledge networks, and it is designed to study equilibrium long-run growth. It is also readily applicable to other networks.

The empirical analysis contributes to the literature on firms' market value and patenting activities, such as Hall et al. (2001, 2005) and Kogan et al. (2017). These papers try to evaluate the economic value of patents and some use citations received as an *ex post* measure of patent quality. For example, Kogan et al. (2017) find positive

[1] An exception is the growth accounting framework by Baqaee et al. (2024) to quantify the importance of supplier churn in firm-to-firm trade for aggregate growth, but such churns are treated as exogenous in their model.

knowledge spillovers and strong, negative creative destruction to peers as in Bloom et al. (2013). In contrast, we find positive impacts on firm value and other financial performances for firms that receive the new pathway citations, because our focus on new pathways across different technology categories limits the negative impact of creative destruction.

Lastly, our definition of pathfinder and path-receiver patents add one more patent quality proxy to the literature summarized by Higham et al. (2021). 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. Our definition of path builder patents are similar to the novel word and the interdisciplinary and cutting-edge measures, in that they are the first patent of its kind in different fashions.

2 Empirics

This section describes the primary data sources and presents the empirical findings on citation network dynamics and their impacts on firm innovation and financial performances.

Data. We use the 2023 version of the U.S. Patent and Trademark Office (USPTO) Patent Database. It contains detailed information on 7.4 million utility patents granted in the U.S. between the years 1976 and 2023. The data also contain each patent's U.S. patent class code to identify its technology category, so that we can construct sectoral level knowledge networks using patent citations. Patenting firms' financial performances matched with Compustat are retrieved from Wharton Research Data Services between 1974 and 2015. 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).

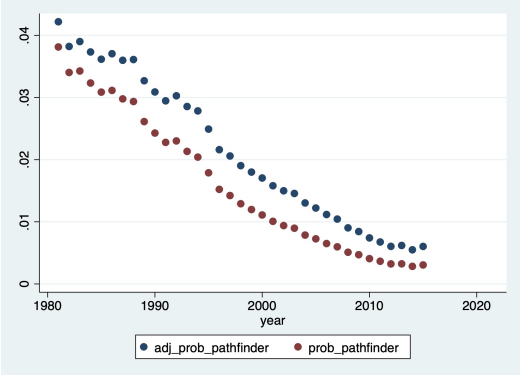


Figure 1: Probability of Observing Pathfinder Patents

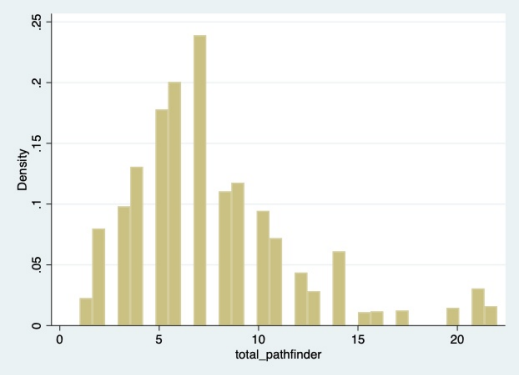


Figure 2: Distribution of Pathfinder Patents across Technology Classes in 2000

Pathways, pathfinders, and path-receivers. We use patents as a proxy of knowledge stock. Patent citations from one technology category to another category indicate the knowledge flow from the cited category to the citing category. A directed citation edge is a pathway from category i to another category j if it is the first citation ever made from i to j. 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. One example is pathfinder patent No. 4110827, US class 700 (Data processing: genera; control systems or specific applications), Load cycling with space temperature feedback, filed in 1976 by pathfinder firm Honeywell Inc., cited path-receiver patent No. 4013118, US class 165 (HEAT EXCHANGE), Control of heating and cooling available from central sources to a multi-zone temperature controlled space, filed in 1975 by path-receiver firm Barber Colman Company. Pathfinder and receiver patents belong to different categories, and hence their owner firms are unlikely to be direct competitors at the product market. Focusing on pathways thus limits the creative destruction effect on the cited firms (path-receivers).

Since this is the first paper on the dynamics formation of citation networks, it is worthwhile to discuss the characteristics of these pathfinder patents. First, pathway formation is a rare event. Fig. 1 shows that pathfinding patents account for around 2% to 3% of all patent applications between 1980 and 2000, and the fraction drops below 1% after 2000. One may suspect that the declining probability of pathfinder

patents is because the availability of unconnected pairs of technology categories is shrinking. However, even after dividing the share of unconnected categories, the adjusted probability (blue dots) still decreases dramatically after 2000. Second, the frequency of observing pathfinder patents has large variation across patent classes in a given year.

Fact 1. *Compared with an average patenting firm, pathfinder and path-receiver firms are much larger in terms of knowledge stock and financial performance measures.*

Table 1: Size Measures of Patenting Firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	employment	marketvalue	sale	profit	tfp	patent stock	no. nclass
pathreceiver	33.57	25157.0	6482.2	2216.7	0.0750	2408.0	22.63
	(50.68)	(31.10)	(39.46)	(39.56)	(10.64)	(76.98)	(111.45)
pathfinder	14.79	10787.5	2165.1	778.4	0.0680	645.9	11.31
	(29.24)	(17.89)	(17.28)	(18.21)	(12.21)	(27.09)	(73.05)
_cons	10.63	3789.3	1830.6	596.2	0.823	116.9	3.997
	(53.92)	(20.05)	(37.74)	(36.02)	(346.57)	(12.67)	(66.70)
N	50413	14542	51720	51680	36683	51796	51796
R^2	0.438	0.368	0.362	0.315	0.123	0.316	0.537

t statistics in parentheses; SIC and year fixed effects included.

Table 1 shows that, compared with average patenting firms, pathfinder and path-receiver firms are much larger in terms of patent stock, the number of patenting technology categories (CPCs), employment, market value, profit, and sales; and slightly higher in TFP.

Fact 2. *Pathways indicate high qualities for both pathfinders and path-receiver patents.*

Results in Table 2 show that pathfinder and path-receiver patents have above-average quality by most quality measures in the literature, except for grant lag and novel word count for path-receivers. Higham et al. (2021)[2] suggest lower backward

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

Table 2: Quality Differences for Path-building Patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1	2	3	4	5	6	7	8	9
pathreceiver	8.169	-0.0978	0.169	0.275	0.0573	0.0672	0.161	4.017	-0.00973
	(2.68)	(-4.52)	(2.08)	(30.04)	(2.46)	(26.22)	(46.37)	(6.81)	(-0.14)
pathfinder	59.33	0.233	0.382	0.305	0.0506	0.211	0.0696	1.003	0.126
	(6.89)	(10.79)	(5.86)	(27.46)	(2.29)	(47.20)	(26.16)	(5.23)	(2.13)
_cons	75.06	2.689	11.88	1.897	1.602	0.520	0.544	8.143	3.267
	(818.90)	(91512.36)	(407325.18)	(20901.65)	(13896.89)	(8128.93)	(7021.83)	(743.68)	(174826.58)
N	1150340	5082950	65456334	4983010	2098476	2459615	1987063	2635185	739918
R^2	0.056	0.330	0.145	0.185	0.135	0.147	0.161	0.146	0.014

t statistics in parentheses. CPC and year fixed effects included. Standard errors clustered by nclass.

1 Citations to scientific literature, 2 Grant Lag, 3 Cited Patents Age, 4 No. of CPCs, 5 Pedigree,

6 Originality, 7 Generality, 8 No. of Forward Citations within 10 years, 9 Novel Word Count.

citations' pedigree, the relative popularity of cited patents, and shorter grant lag, the time difference between grant date and application data, predict lower patent quality. Balsmeier et al. (2018) use weighted novel word count to evaluate a patent's originality. Pathfinders are more likely to be *new combinations* of existing ideas from different technology categories, such as finding a new use of an existing product. Hence, they may not create a brand new idea as patents with novel words do.

Impacts of knowledge network dynamics on firms. We then examine the impact of citation network dynamics on firms' innovation and market performance.[3] Our first task is to construct the measures of firms' involvement in knowledge pathway creations in four different ways. A pathway from category i to category j directly connects the pathfinder and receiver firms. A pathway event also establishes the public information that knowledge in j is useful in producing new knowledge (patents) in i through the patent system. Therefore, such an event may have spillover effects on peer firms who also innovate in i and j. We are interested in both the direct impacts on pathfinder and receiver firms and the indirect spillovers to their peers.

[3]Some researchers worry that the knowledge network pathway events could be driven by other firm level shocks such as demand shocks. In fact, the innovation process takes much longer time and bares much larger uncertainty than production, which provides a years-long and random gap between the arrival of demand shocks and the filing/publication of new citation pathways in the granted patent applications, therefore it is valid to assume the independence of knowledge pathway events and firms' other shocks.

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 in citation network dynamics: 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 $\text{wpf}_{f,t} = \sum_j w^f_{j,t} pf^f_{j,t}$ proxy firm f's direct involvement in pathway creation; weighted number of pathways received and found by f's peers, $\text{wpr}_{-f,t} = \sum_j w^f_{j,t} pr^{-f}_{j,t}$ and $\text{wpf}_{-f,t} = \sum_j w^f_{j,t} pf^{-f}_{j,t}$, indicate firm f's indirect closeness to new pathways built by peers; where $pr^f_{j,t}$, $pf^f_{j,t}$, $pr^{-f}_{j,t}$ and $pf^{-f}_{j,t}$ are the numbers of pathways received and found by firm f itself and peer firms $-f$ in year t category j, respectively. Since each firm may have patent stocks in many categories, the weights are f's patent stock shares across categories, such that $w^f_{j,t} = N^f_{j,t}/N^f_t$, $N^f_{j,t}$ 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.

Equipped with firms' exposure measures to network dynamics, we then use local projection method in Jordà (2005) to estimate the four pathway creation measures' impact on firm performances, such as granted patent stock, sales, profit, employment, capital, market value, and TFP over a horizon of one to five years after the pathways' appearance[4]. The regressions belong to the following format.

$$\begin{aligned}\log\left(X_{f,t+s}\right) =& \underline{\mathrm{p}}^s_1 path_{f,t-1} + ... + \underline{\mathrm{p}}^s_p path_{f,t-p} + \underline{\mathrm{x}}^s_1 \log(X_{f,t-1}) + ... + \underline{\mathrm{x}}^s_p \log(X_{f,t-p}) \\ &+ \underline{\mathrm{z}}^s_1 Z_{f,t-1} + ... + \underline{\mathrm{z}}^s_p Z_{f,t-p} + u_{f,t+s}, \quad s = 0, 1, 2, 3, 4,\end{aligned} \tag{1}$$

where the vector $path_{f,t}$=$(\text{lwpf}_{f,t}, \text{lwpr}_{f,t}, \text{lwpf}_{-f,t}, \text{lwpr}_{-f,t})$ measures each firm's direct and indirect involvement in pathway creations in year t. $X_{f,t}$ is a performance 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 $Z_{f,t}$ represents a set of controls including firm, industry (SIC4) and year fixed effects, together with other variables that differ across regressions. In Table 2, we know that pathways indicate high qualities of both the finder and the receiver patents. We expect firms to benefit from either owning high quality patents, sourcing/providing

[4]We do not intend to establish a causality relation between a pathway creation and self firm performances, because the network formation is an endogenous decision.

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

knowledge from/to new categories in the future, or observing other firms paving new knowledge pathways, that is, the vector $\underline{p}^s_q$ to be positive for $s = 0, 1, 2, 3, 4$ and $0 \leq q \leq p$.

I. Pathways and firms' innovation activities. Let $X_{f,t}$ be firm f's patent stock at t. Fig. 3 provides the result when the dependent variable in Eq. (1) is the firm's patent stock over a five-year horizon. Controls in $Z_{f,t}$ include firm f's number of technology categories in which firm f has granted patent applications at $t-1$,..., $t-p$, and firm and year fixed effects. As expected, patent stock of all firms that involved in pathway creation increases in the future 5 years, reflected by the positive and significant $\underline{p}^s_{t-q}$, $s = 0, 1, 2, 3, 4$ and $0 \leq q \leq p = 3$. 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. [6]

Fact 3. *Patent stocks of pathfinder and receiver firms and their peers increase faster after pathways appear.*

The coefficients capturing the spillover effects on peers in Fig. 3 appear to be magnitudes larger than coefficients of direct impacts. The reasons are twofold. One is that, as Table 1 shows, pathfinder and receiver firms tend to be much larger than their peers and thus less sensitive on the growth margin. In robustness checks 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.

II. Pathways and firm performances. We turn to firms' product-market performance measures and focus on publicly traded firms in Compustat only. In this case, $X_{f,t}$ in Eq. (1) can be a firm's profit, sales, market value, total employment, capital stock, or productivity. Controls in $Z_{f,t}$ include two measures of a firm's

[6] We cannot control for the industry fixed effects for all patenting firms in the regression for Fig. 3, because industry identity at product market is only available to public firms in Compustat, and most firms apply patents in multiple patent technology categories, therefore a single patent category is not a good industry identity for a firm.

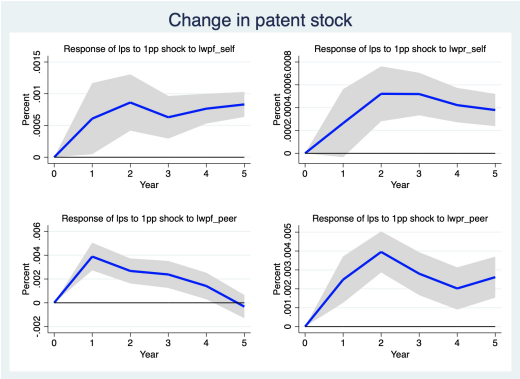


Figure 3: Knowledge Path Creation and Firm Patent Stock

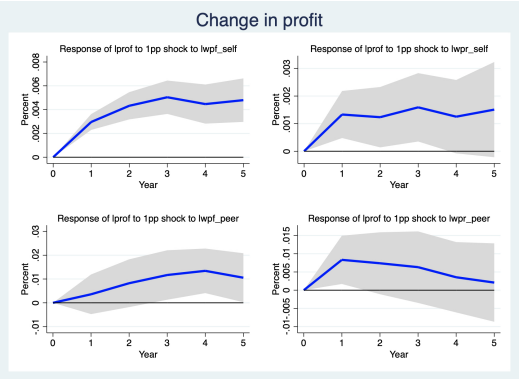


Figure 4: Knowledge Path Creation and Firm Profit

innovation output in logarithms following Kogan et al. (2017): a stock-market based measure of patents' dollar values and the citation-weighted patent stock. Here we set $p = 1$ because of the limited number of observations for publicly listed firms. Firm industry (SIC4) and year fixed effects remain in place.

Fact 4. *Public pathfinder and receiver firms and their peers grow faster in sales, profit, employment, capital, market value and TFP after pathways appear.*

Fig. 4 presents the local projection results on profit growths, more results with other financial performances and regression tables are provided in Appendix D.1. Consistent with intuitions and previous studies, the results suggest that a pathfinder firm directly benefits from owning a new, high-quality pathway patent, and the possibility of absorbing knowledge input from a new source in the future. 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. Akin to the finding in Fig. 3, the impact on peer firms are often stronger than on path building firms themselves. A path-receiver firm benefits, with coefficient values comparable to a pathfinder's or even more prominent. By definition, pathfinder patents have limited creative destruction effects on those in the receiving category, because they belong to different technology categories. Since all patents are somewhat substitutable in the pathway receiving category, the newly paved knowledge pathway brings attention and potential new uses to all patents in the cited category, all peer firms in the knowledge used

category benefit through future forward citations, royalty payment and reputation.

Robustness checks. A set of robustness checks using different measures of knowledge pathway creation are provided in Appendix D.2. We define the new knowledge pathways using the first 10 and 100 citations across a category pair to consider the situation when many inventors discover the same new knowledge path at roughly the same time; and we define the new pathways as the first pathways within the last 10 years to consider the cases when some knowledge paths are obsolete and then reused after more than a decade. Our findings are robust to these alternative measures of pathways.

Takeaways. The paper aims to study the growth implications of knowledge network dynamics. Therefore, we need a dynamic model with the following elements. First and foremost, new pathways are rarely observed compared to other citations, so the model must have frictions in new-link formation. Second, as Facts 1 and 2 indicate, new connections across categories likely appear between large firms and high-quality innovations, so the heterogeneity in firm innovation outcome and the likelihood of new-link formation in the model must be related. Third, as Facts 3 and 4 suggest, a new pathway benefits both the pathfinder and the receiver, and such benefits spill over to their peers, so the model must feature such positive spillovers in firms' payoffs.

3 Model

We begin with a simplified environment to highlight the endogenous dynamics *in* and *of* the knowledge network, without modeling the spillovers among firms. Then, the extended model in Section 5 allows for richer heterogeneity and explicitly considers the spillover effect when new links are formed in the network. Throughout, we focus on equilibria with time-invariant firm value function and policies, featuring aggregate stationarity and idiosyncratic dynamism.

3.1 Setup

Time is continuous and infinite, $t \in \mathbb{R}$. The economy has a fixed continuum of industries with unit measure. An industry is a technological category with a distinct piece of knowledge that can be used as an innovation input. In this simplified setup,

each industry has one firm. Each firm receives profits from the product market and discounts future cashflows at a common rate $\rho > 0$. Firms hire researchers at a constant wage rate $w > 0$ from a *frictionless* market. Forward-looking firms maximize the present value of expected cashflows. Information is complete. Firms face search frictions when forming new links with each other in the network.

Each firm has a fixed two-dimensional *innovation efficiency type* denoted as $(\tau, \epsilon) \in [\tau_0, \tau_1] \times [\epsilon_0, \epsilon_1] \subseteq \mathbb{R}^2_+$, where τ is referred to as the firm's *output* type, governing the firm's own innovation outcome, whereas ϵ is its *input* type, affecting other firms' innovation outcome whenever it serves as a supplier of knowledge input. For example, when firm i is a supplier to firm j, i.e., j is a user of input i, firm j's innovation outcome depends on τ_j and ϵ_i. The output type τ can be thought of as a firm's efficiency in using others' knowledge to innovate; the input type ϵ can be considered as the quality, applicability, or some "upstreamness" measure of a firm's knowledge as an input. These types are distributed across firms according to a cumulative distribution function (cdf) F with marginal cdf's respectively denoted as F_τ and F_ϵ and respective means $\overline{\tau}$ and $\overline{\epsilon}$. We keep firm types fixed to highlight network evolution as the only force driving idiosyncratic dynamics. Note that a firm's τ and ϵ need not be independent; for example, a high-ϵ firm may also have a high τ.

Remark (Notation). We use the following notational rule henceforth. For a generic variable or function x, hat $\widehat{x}$ indicates an individual firm's *choice*, tilde $\widetilde{x}$ reflects *uncertainty*, bar $\overline{x}$ is the cross-sectional *mean*, bold $\boldsymbol{x}$ represents an *equilibrium or aggregate* object that an individual firm must take as given.

Each firm's profit flow depends on its productivity level *relative* to the cross-sectional average at the moment. Specifically, if a firm's productivity level is $Z > 0$ at t when the *average* productivity level is $\boldsymbol{Z}(t) > 0$, the firm receives an instantaneous profit of πz, where $\pi > 0$ is the aggregate profit flow, exogenous for now, and $z := Z/\boldsymbol{Z}(t)$ is the firm's *relative productivity*. By construction, the cross-sectional average of the *relative* productivity z is always 1. Here, $\boldsymbol{Z}(t)$ aggregates the productivity levels of all firms, and its constant growth rate $\boldsymbol{g} := \dot{\boldsymbol{Z}}(t)/\boldsymbol{Z}(t) \geq 0$ is a major object of interest at the stationary equilibrium. From an individual firm's viewpoint, it is as if its relative productivity z depreciates at rate $\boldsymbol{g}$. Therefore, if the firm's productivity *level* remains at Z from t onwards absent all future innovations, its expected profit

stream valued at t can be written as zv_z, such that

$$\rho z v_z = \pi z - \boldsymbol{g} z v_z, \tag{2}$$

so the constant marginal value of relative productivity is $v_z = \frac{\pi}{\rho + \boldsymbol{g}} > 0$, endogenously determined in equilibrium and common to all firms. It will become clear that v_z gives firms the incentives to innovate to raise Z and z.

Innovation requires knowledge inputs. Every firm, representing an industry, is a potential supplier of knowledge input to others and a potential user of others' knowledge. A firm must search for input suppliers before being able to use them for innovation; it must be found by others to become their supplier. These supplier-user relationships form the *directed knowledge network* with the continuum of firms as nodes and each supplier-user relationship as a directed link. Link direction is from a user to a supplier, indicating the (inverse) direction of knowledge flow.

In what follows, consider a type-(τ, ϵ) firm at t that has relative productivity z and has accumulated a set U of other firms as upstream suppliers and a set D of other firms as downstream users. Each supplier $u \in U$ corresponds to an available knowledge input u. It is understood that (z, U, D) evolves over time while (τ, ϵ) is fixed. The supplier count $|U| \in \mathbb{N}$ and user count $|D| \in \mathbb{N}$ respectively correspond to the numbers of *pathways found* (outdegree) and *received* (indegree) at t. Let $(\tau, \epsilon, z, U, D)$ summarize the firm's state and $V(\tau, \epsilon, z, U, D)$ a firm's equilibrium value at this state. It will be shown that firm value has an additively separable form $V(\tau, \epsilon, z, U, D) = zv_z + v(\tau, \epsilon, U, D)$, where zv_z is the value of the relative productivity z at the moment, satisfying Eq. (2), and $v(\tau, \epsilon, U, D)$ is the continuation value encompassing the firm's decisions to innovate and to form new links.

Dynamics in the existing network: innovation and surplus sharing. Innovation entails hiring researchers and using available knowledge inputs to raise productivity with random success. Specifically, if the firm wants to achieve a Poisson success rate $\widehat{\mu}$ while exploiting U, it must hire $\tau \epsilon^U c_\mu(\widehat{\mu})$ units of research labor, where $c_\mu(\mu)$ satisfies $c_\mu(0) = 0$, $c'_\mu(\mu), c''_\mu(\mu) > 0$ for $\mu > 0$, and $\epsilon^U := \sum_{u \in U} \epsilon_u$ is the combined input type of all suppliers in U, i.e., the type-weighted supplier count. Using more and better inputs to innovate costs proportionally more researchers, which will ensure a scale-free decision of innovation success rate as in Klette and Kortum (2004).

The chosen $\widehat{\mu}$ governs the frequency of successful innovations. Conditional on

a success, the firm's productivity *level* Z jumps up by $\tau\epsilon^U \boldsymbol{Z}(t)$; the states of the suppliers in U are unchanged. The average productivity $\boldsymbol{Z}(t)$ in the jump size reflects the usual *knowledge level externality* in innovation and ensures stationary aggregate growth. Equivalently, the jump in relative productivity is simply $\tau\epsilon^U$, so the relative productivity z evolves as $dz = -\boldsymbol{g}zdt + \tau\epsilon^U d\mathrm{P}(\widehat{\mu}, t)$ given U at t, where $\mathrm{P}(\widehat{\mu}, t)$ represents a Poisson process at t with parameter $\widehat{\mu}$. The pairwise contribution $\tau\epsilon_u$ to the jump is referred to as the *match quality* between the firm as a user and its supplier $\forall u \in U$. Having a higher output type and using more and better knowledge inputs thus lead to greater innovation outcomes. Note that firms always use all inputs in U even if they can choose to use any subset of it.

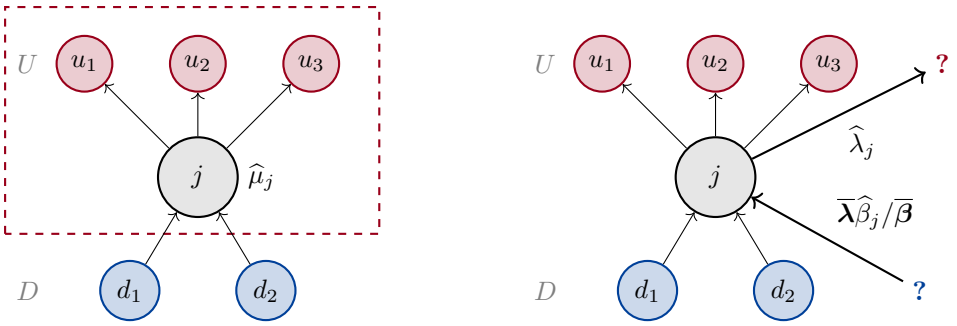
A jump in productivity raises the innovator's expected profit stream, and the value of such raises is the *innovation surplus*. With an additively separable firm value, let $S(\tau, \epsilon^U) := V(\tau, \epsilon, z + \tau\epsilon^U, U, D) - V(\tau, \epsilon, z, U, D) = \tau\epsilon^U v_z$ denote the endogenous innovation surplus using U. Note the surplus separability by suppliers, $S(\tau, \epsilon^U) = \sum_{u\in U} S(\tau, \epsilon_u)$. The surplus is perfectly transferable and split multiway among the innovator firm and its suppliers in U as follows. Each supplier $u \in U$ receives $bS(\tau, \epsilon_u)$, where parameter $b \in (0, 1)$ is the supplier's share in each supplier-user match; the innovator firm gets all of the remaining $(1-b)S(\tau, \epsilon^U)$.[7] The supplier surplus $bS(\tau, \epsilon_u)$ may be viewed as a transfer from the innovator to each supplier u. At the special case $b = \frac{1}{2}$, the surplus allocation rule is the standard Shapley value solution.[8] As $b \to 0$, the innovator gains more from innovation, and suppliers extract lower rents when users innovate. The cost of innovation, i.e., researcher wage bill, is already sunk at the stage of surplus sharing, so the innovator bears all of it. The innovator's expected net flow payoff at a chosen innovation rate $\widehat{\mu} \geq 0$ is

$$\widehat{\mu}(1-b)S(\tau, \epsilon^U) - w\tau\epsilon^U c_\mu(\widehat{\mu}). \tag{3}$$

Recall that the focal firm has input type ϵ and is a supplier to all firms in its

[7] The surplus allocation rule is consistent with setting a royalty rate in licensing agreements as a percentage of sales of relevant products, as is common in practice. Here, supplier u gets a fraction $b\epsilon_u/\epsilon^U$ of the expected value of future sales $\tau\epsilon^U v_z$ as a result of using said supplier u's input.

[8] Thanks to each supplier's separable contribution to innovation surplus, the surplus allocation rule also coincides with symmetric pairwise Nash bargaining. In the extended model Section 5, the standard Shapley value solution corresponds to a $b \neq \frac{1}{2}$ because of a within-industry supplier switching cost.



(a) firm j innovates at rate $\widehat{\mu}_j$ using inputs from $U = \{u_1, u_2, u_3\}$ and shares surplus with them

(b) firm j expects to get a new supplier at $\widehat{\lambda}_j$ and a new user at $\overline{\boldsymbol{\lambda}}\widehat{\beta}_j/\overline{\boldsymbol{\beta}}$

Figure 5: Dynamics in and of the knowledge network: an illustration from firm j's perspective.

downstream set D. Hence, a successful downstream innovation by user $d \in D$ arrives at its chosen rate μ_d and brings a revenue flow $bS(\tau_d, \epsilon) = b\tau_d \epsilon v_z$, where τ_d is the user's output type, and user d has v_z as the marginal value of relative productivity. Note that the rent from user d does not depend on d's other suppliers. The firm's expected flow rent extracted from all users in D is

$$b \sum\nolimits_{d \in D} \mu_d S(\tau_d, \epsilon), \tag{4}$$

without any costs. Having a higher input type and supplying to more users of higher output types thus raises a firm's revenue flow.

The innovation process and surplus allocation have a convenient consequence. To a firm of type (τ, ϵ), each of its supplier $u \in U$ is payoff-relevant only via u's input type ϵ_u and the firm's own output type τ, whereas each user $d \in D$ is payoff-relevant only via d's innovation rate μ_d and output type τ_d, and the firm's own input type ϵ. It is thus sufficient to have $U = \{\epsilon_u\}_{u \in U}$ and $D = \{(\tau_d, \mu_d)\}_{d \in D}$ in firm state with some notational abuse of U and D.

Successful innovation is the only way for firms to raise productivity. Hence, the aggregate growth $\boldsymbol{g}$ in $\boldsymbol{Z}(t)$ is the cross-sectional aggregation of all firms' $\mu\tau\epsilon^U$, where μ is each firm's chosen rate of successful innovations, τ is own output type, and ϵ^U is its available suppliers' combined input type, i.e., type-weighted supplier count. Each firm's upstream set U, however, keeps evolving endogenously, so $\boldsymbol{g}$ depends on the network dynamics via each firm's ϵ^U.

Dynamics of the network: link removal and addition. We build on the Diamond-Mortensen-Pissarides (DMP) framework of random search and matching to

model network evolution with endogenous formation of new links and random removal of old ones. The exogenous Poisson rate of link destruction is $\delta > 0$, i.e., the average time that a supplier's knowledge serves as an applicable input for a user is $1/\delta$. It reflects a form of knowledge obsolescence that is specific to each supplier-user pair: input i may still be useful to some user firm j when firm $j' \neq j$ ceases to use it.

Formation of new links is driven by each firm *searching* for new suppliers and *attracting* new users simultaneously.[9] The meeting between a searching firm (a potential new user) and an attracting firm (a potential new supplier) is random but depends on the endogenous efforts on both sides. Each firm chooses a search intensity $\widehat{\lambda}$ and an attraction level $\widehat{\beta}$ at the search cost of $wc_\lambda(\widehat{\lambda})$ and the attraction cost of $wc_\beta(\widehat{\beta})$, both as researcher wage bills, such that for $\iota = \lambda, \beta$, $c_\iota(s)$ satisfies $c_\iota(0) = 0$ and $c_\iota'(s), c_\iota''(s) > 0$ for $s > 0$. The total mass of meetings at any time is $\overline{\boldsymbol{\lambda}}$, the cross-sectional aggregation of individual search intensity, allocated within each side proportional to the corresponding efforts. On the searching side, meetings are allocated by search intensities, such that a firm that chooses to search at $\widehat{\lambda}$ meets a new supplier on the attracting side at Poisson rate $\widehat{\lambda}$. On the attracting side, meetings are allocated by attraction levels, such that a firm with attraction level $\widehat{\beta}$ meets a new user on the searching side at Poisson rate $\overline{\boldsymbol{\lambda}}\widehat{\beta}/\overline{\boldsymbol{\beta}}$, where $\overline{\boldsymbol{\beta}}$ denotes the cross-sectional aggregation of attraction levels.

Such a matching technology does not have search externality on the searching side, which reflects *knowledge non-rivalry*. Any piece of knowledge can be used repeatedly as an innovation input by one or multiple users. Therefore, one firm using a particular input or searching for new inputs does not affect others' rate of finding them. In contrast, congestion occurs on the attracting side as suppliers compete for user attention, such that "better" inputs are more likely used despite random matching. Endogenous attraction can be considered as firms investing in the generality or applicability of their knowledge.

A new supplier-user link, i.e., a pathway, is formed when two firms meet. Indeed, without link formation cost or capacity constraints, firms always benefit from forming new links. To a firm with (τ, ϵ, U, D), the value of a new supplier with input type $\widetilde{\epsilon}$ is

[9] We distinguish "searching" and "attracting" to clarify the directions of each firm's two-way searching behavior, such that "searching" is to search for upstream suppliers, whereas "attracting" is to search for downstream users.

$v(\tau, \epsilon, U \cup \{\widetilde{\epsilon}\}, D) - v(\tau, \epsilon, U, D) \geq 0$. The firm's expected flow payoff when searching at $\widehat{\lambda}$ is

$$\widehat{\lambda}\, \mathbb{E}^{\widetilde{\beta},\widetilde{\epsilon}}\left[\frac{\widetilde{\beta}}{\overline{\boldsymbol{\beta}}}[v(\tau, \epsilon, U \cup \{\widetilde{\epsilon}\}, D) - v(\tau, \epsilon, U, D)]\right] - wc_\lambda(\widehat{\lambda}), \tag{5}$$

where $\mathbb{E}^{\widetilde{\beta},\widetilde{\epsilon}}[\cdot]$ is the expectation operator defined with the joint cross-sectional distribution of all firms' attraction choices $\widetilde{\beta}$ and input types $\widetilde{\epsilon}$. Similarly, the value of a new user with output type $\widetilde{\tau}$ and innovation rate $\widetilde{\mu}$ is $v(\tau, \epsilon, U, D \cup \{(\widetilde{\tau}, \widetilde{\mu})\}) - v(\tau, \epsilon, U, D) \geq 0$. The firm's expected payoff at attraction level $\widehat{\beta}$ is

$$\frac{\widehat{\beta}}{\overline{\boldsymbol{\beta}}}\, \mathbb{E}^{\widetilde{\lambda},\widetilde{\mu},\widetilde{\tau}}\left[\widetilde{\lambda}[v(\tau, \epsilon, U, D \cup \{(\widetilde{\tau}, \widetilde{\mu})\}) - v(\tau, \epsilon, U, D)]\right] - wc_\beta(\widehat{\beta}), \tag{6}$$

where $\mathbb{E}^{\widetilde{\lambda},\widetilde{\mu},\widetilde{\tau}}[\cdot]$ is the expectation operator defined with the joint cross-sectional distribution of all firms' search intensities $\widetilde{\lambda}$, innovation rates $\widetilde{\mu}$, and out types $\widetilde{\tau}$. When a firm chooses the search intensity $\widehat{\lambda}$ and the attraction level $\widehat{\beta}$, it considers all other firms' searching, attraction, and innovation decisions and trades off the search and attraction costs for gradual accumulation of new suppliers and new users.

For any firm, each existing supplier $u \in U$ or user $d \in D$ also gains or loses their suppliers (U_u or U_d) and users (D_u or D_d) over time, and so do all the firms *indirectly* linked to it. In general, these events should affect a firm's value as well. Here, these indirect connections and their changes have no impact on a firm's value, which drastically simplifies the analysis. See Appendix A.2 for further discussions.

Summary of firm problem. See Fig. 5 for an illustration. At any time t, a firm chooses an innovation rate $\widehat{\mu}$ using U and decides on a search intensity $\widehat{\lambda}$ and an attraction level $\widehat{\beta}$ to expand U and D, facing respective costs in researcher hiring. The innovating firm and all of its suppliers share the surplus at each success. The firm makes decisions based on its own type (τ, ϵ), relative productivity z, and directly linked firms in U and D. The firm also takes as given the joint distribution of all firms' states and actions and an aggregate growth rate $\boldsymbol{g}$.

3.2 Firm value and policies

We have the main pieces Eq. (2) to Eq. (6) of the Hamilton-Jacobi-Bellman (HJB) equation for a particular firm's value $V(\tau, \epsilon, z, U, D) = zv_z + v(\tau, \epsilon, U, D)$. Notice that the flow payoffs associated with the supplier set U are independent of those

concerning the user set D. Hence, $v(\cdot)$ can be written in two separate parts, $v^{up}(\cdot)$ with upstream origins and $v^{down}(\cdot)$ with downstream origins.

Suppose that the particular firm anticipates all the others to innovate at a common and time-invariant rate $\boldsymbol{\mu} > 0$. Then, existing users in D are payoff-relevant only via their output types, and D simplifies to $D = \{\tau_d\}_{d \in D}$. Let Ω be the set of all possible realizations of firm states. Suppose that the firm also takes as given a stationary cross-sectional distribution $\boldsymbol{M}$ over Ω and other firms' search and attraction policies as $\boldsymbol{\lambda}, \boldsymbol{\beta} : \Omega \to \mathbb{R}_+$, respectively. The average attraction level follows as $\overline{\boldsymbol{\beta}} = \int \boldsymbol{\beta}(\widetilde{\omega}) d\boldsymbol{M}$. Moreover, the firm anticipates the aggregate growth rate to be $\boldsymbol{g} = \boldsymbol{\mu} \int \widetilde{\tau} \widetilde{\epsilon^U} d\boldsymbol{M}$. Therefore, at the given $(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{M})$, we write firm value as

$$V(\tau, \epsilon, z, U, D) = z v_z + v^{up}(\tau, U) + v^{down}(\epsilon, D), \tag{7}$$

where the HJB equation for $z v_z$ is as given in Eq. (2), and the respective HJB equations for $v^{up}(\cdot)$ and $v^{down}(\cdot)$ given the surplus allocation rule are

$$\begin{aligned}
\rho v^{up}(\tau, U) = \max_{\widehat{\mu}, \widehat{\lambda}} \Bigg\{ & \widehat{\mu}(1-b) S(\tau, \epsilon^U) - w \tau \epsilon^U c_\mu(\widehat{\mu}) + \delta \sum\nolimits_{u \in U} [v^{up}(\tau, U \setminus \{\epsilon_u\}) - v^{up}(\tau, U)] \\
& + \widehat{\lambda} \int \frac{\boldsymbol{\beta}(\widetilde{\omega})}{\overline{\boldsymbol{\beta}}} [v^{up}(\tau, U \cup \{\widetilde{\epsilon}\}) - v^{up}(\tau, U)] d\boldsymbol{M} - w c_\lambda(\widehat{\lambda}) \Bigg\}, \qquad (8) \\
\rho v^{down}(\epsilon, D) = \max_{\widehat{\beta}} \Bigg\{ & b \sum\nolimits_{d \in D} \boldsymbol{\mu} S(\tau_d, \epsilon) + \delta \sum\nolimits_{d \in D} [v^{down}(\epsilon, D \setminus \{\tau_d\}) - v^{down}(\epsilon, D)] \\
& + \frac{\widehat{\beta}}{\overline{\boldsymbol{\beta}}} \int \boldsymbol{\lambda}(\widetilde{\omega}) [v^{down}(\epsilon, D \cup \{\widetilde{\tau}\}) - v^{down}(\epsilon, D)] d\boldsymbol{M} - w c_\beta(\widehat{\beta}) \Bigg\}. \qquad (9)
\end{aligned}$$

The firm's value change due to the productivity jump at a successful innovation is embedded in Eq. (8) as the expected innovation surplus $\widehat{\mu} S(\tau, \epsilon^U)$, and a fraction b goes to its suppliers; rents to suppliers are mirrored in Eq. (9) as rents received from users. Eqs. (8) and (9) both include the consequence of exogenous link destruction and the endogenous decisions to form new upstream or downstream links. A firm's value obviously depends on other firms' decisions summarized by the list $(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{M})$ and the inferred $\boldsymbol{g}$, dropped from the expressions to ease notation. Similarly, it is understood that the associated optimal policies denoted as $\mu^*(\cdot)$, $\lambda^*(\cdot)$, and $\beta^*(\cdot)$ are the firm's best response to others' decisions.

4 Equilibrium Analysis and Implications

This section characterizes the stationary equilibrium analytically. All proofs are in Appendix A. A few modeling assumptions are discussed in Section 4.5.

Definition 1. A *stationary equilibrium* consists of a common innovation rate $\boldsymbol{\mu} > 0$, search and attraction policies $\boldsymbol{\lambda}, \boldsymbol{\beta} : \Omega \to \mathbb{R}_+$, and a cross-sectional state distribution $\boldsymbol{M}$ over Ω, an aggregate growth rate $\boldsymbol{g} \geq 0$, and firm value $V : \Omega \to \mathbb{R}_+$, such that given $(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{M})$, each firm with $\forall \omega = (\tau, \epsilon, z, U, D) \in \Omega$ infers the aggregate growth rate as $\boldsymbol{g} = \boldsymbol{\mu} \int \widetilde{\tau}\widetilde{\epsilon^U} d\boldsymbol{M}$ and optimizes to achieve value $V(\omega)$ with policies $\mu^*(\omega) = \boldsymbol{\mu}$, $\lambda^*(\omega) = \boldsymbol{\lambda}(\omega)$, $\beta^*(\omega) = \boldsymbol{\beta}(\omega)$, and $\boldsymbol{M}$ is induced by the type distribution F, policies $(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta})$, and exogenous link destruction at δ.

4.1 Characterization of the equilibrium

To characterize the equilibrium, we need to know how the cross-sectional distribution $\boldsymbol{M}$ of all firms enters each individual firm's problem directly in Eqs. (8) and (9) and indirectly as growth rate $\boldsymbol{g}$ in Eq. (2). Crucially, $\boldsymbol{M}$ depends on firms' search and attraction decisions to form the network. Fortunately, this simple setup allows for a few sufficient statistics to replace $\boldsymbol{M}$ as an aggregate state variable.

Suppose that the equilibrium search and attraction policies are type-specific, such that $\boldsymbol{\lambda} : [\tau_0, \tau_1] \to \mathbb{R}_+$ and $\boldsymbol{\beta} : [\epsilon_0, \epsilon_1] \to \mathbb{R}_+$ are strictly increasing functions. We will show that it is indeed the case in equilibrium. Then, the aggregate search rate $\overline{\boldsymbol{\lambda}}$ and the aggregate attraction level $\overline{\boldsymbol{\beta}}$ become

$$\overline{\boldsymbol{\lambda}} = \int \boldsymbol{\lambda}(\tau) dF_\tau(\tau) > 0, \quad \text{and} \quad \overline{\boldsymbol{\beta}} = \int \boldsymbol{\beta}(\epsilon) dF_\epsilon(\epsilon) > 0, \tag{10}$$

where F_τ and F_ϵ are model primitives for marginal type cdfs across firms, i.e., *nodes*.

Alternatively, we can examine the type distribution across all matched supplier-user pairs, i.e., *links*. Each randomly drawn link connects a user firm of some output type τ with a supplier firm of some input type ϵ', i.e., the pair has match quality $\tau\epsilon'$. With random matching and random separation, τ and ϵ' are independent, and the joint cdf of this type duo denoted as $H(\tau \to \epsilon')$ satisfies

$$H(\tau \to \epsilon') = \frac{\int_{\tau_0}^{\tau} \boldsymbol{\lambda}(\widetilde{\tau}) dF_\tau(\widetilde{\tau})}{\overline{\boldsymbol{\lambda}}} \times \frac{\int_{\epsilon_0}^{\epsilon'} \boldsymbol{\beta}(\widetilde{\epsilon}) dF_\epsilon(\widetilde{\epsilon})}{\overline{\boldsymbol{\beta}}} < F_\tau(\tau) \times F_\epsilon(\epsilon'). \tag{11}$$

It means that a link more likely appears between a better user with greater τ and a better supplier with greater ϵ'. Therefore, the knowledge network can be viewed as a weighted one with each link's *match quality* being its *weight*. Moreover, despite the exogenous node types, these weights have an *endogenous* distribution due to endogenous link formation. The mean match quality conditional on being matched is thus $\int \tau\epsilon' dH(\tau \to \epsilon') = \overline{\boldsymbol{\tau}_d} \times \overline{\boldsymbol{\epsilon}_u}$, such that

$$\overline{\boldsymbol{\tau}_d} := \frac{\int \tau \boldsymbol{\lambda}(\tau) dF_\tau(\tau)}{\overline{\boldsymbol{\lambda}}} > \overline{\tau}, \quad \text{and} \quad \overline{\boldsymbol{\epsilon}_u} := \frac{\int \epsilon \boldsymbol{\beta}(\epsilon) dF_\epsilon(\epsilon)}{\overline{\boldsymbol{\beta}}} > \bar{\epsilon}, \tag{12}$$

where $\overline{\tau}$ and $\bar{\epsilon}$ are the unconditional means of respective types. We can interpret $\overline{\boldsymbol{\tau}_d}$ as the mean output type of active users and $\overline{\boldsymbol{\epsilon}_u}$ the mean type of inputs in use. Although the model abstract from extensive-margin selections, the conditional means of types are higher than the unconditional ones due to endogenous type-dependent efforts.

In this simple setup, $\overline{\boldsymbol{\lambda}}$, $\overline{\boldsymbol{\beta}}$, $\overline{\boldsymbol{\tau}_d}$, and $\overline{\boldsymbol{\epsilon}_u}$ are sufficient to capture how $\boldsymbol{M}$ enters firm decisions. We characterize the equilibrium by addressing three major questions.

4.1.1 How do knowledge network dynamics drive aggregate growth?

Proposition 1. *The aggregate growth rate in equilibrium satisfies*

$$\boldsymbol{g} = \boldsymbol{\mu}\boldsymbol{X} = \underbrace{\boldsymbol{\mu}}_{\text{user innov.}} \times \underbrace{\overline{\boldsymbol{\tau}_d}}_{\text{user type}} \times \underbrace{\boldsymbol{x}}_{\text{\# links}} \times \underbrace{\overline{\boldsymbol{\epsilon}_u}}_{\text{supplier type}}, \tag{13}$$

where $\overline{\boldsymbol{\lambda}}$ *is in Eq. (10),* $\overline{\boldsymbol{\tau}_d}$ *and* $\overline{\boldsymbol{\epsilon}_u}$ *are in Eq. (12),* $\boldsymbol{x} = \overline{\boldsymbol{\lambda}}/\delta$ *is the* (unweighted) average degree *of the knowledge network, and* $\boldsymbol{X} = \int \tau\epsilon^U d\boldsymbol{M} = \int \tau^D \epsilon d\boldsymbol{M} = \overline{\boldsymbol{\tau}_d}\boldsymbol{x}\overline{\boldsymbol{\epsilon}_u} > \overline{\tau}\boldsymbol{x}\bar{\epsilon}$ *is the* match-quality-weighted average degree.

Proposition 1 highlights the knowledge-network origins of aggregate growth. Eq. (13) explicitly connects $\boldsymbol{g}$ with knowledge network *density* and *quality*. Aggregate productivity grows because firms innovate at $\boldsymbol{\mu}$ using existing suppliers previously found. Faster growth is associated with a denser knowledge network with more links $\boldsymbol{x}$. The average degree $\boldsymbol{x}$ of a directed network is the average number of links that a node has in either direction, $\boldsymbol{x} = \overline{x^U} = \overline{x^D}$, where $x^U := |U|$ is a firm's supplier count (outdegree) and $x^D := |D|$ the user count (indegree). The average in- and outdegrees are always equal in any directed network as an outward link of one node must be an inward link of another node. The network is denser with more links across firms

when firms meet more frequently. Faster growth is also associated with higher *match qualities* averaged across links at $\overline{\boldsymbol{\tau}_d} \times \overline{\boldsymbol{\epsilon}_u}$, also governed by firm search and attraction policies. The match-quality-weighted average degree $\boldsymbol{X}$ aggregates match qualities across all links and encapsulates the measures of *density* $\boldsymbol{x}$ and *quality* $\overline{\boldsymbol{\tau}_d} \times \overline{\boldsymbol{\epsilon}_u}$.

The extended model in Section 5 allows for non-network-based growth as well and yields a richer version of Eq. (13) in Eq. (20). Nonetheless, the essential idea remains that knowledge network dynamics are crucial for growth.

4.1.2 What are the implications on firm value and decisions?

Proposition 2. *In a stationary equilibrium, firm state simplifies to* $(\tau, \epsilon, z, \epsilon^U, \tau^D)$, *where* $\epsilon^U = \sum_{u \in U} \epsilon_u$ *and* $\tau^D = \sum_{d \in D} \tau_d$. *Firm value function has the form*

$$V(\tau, \epsilon, z, \epsilon^U, \tau^D) = \underbrace{z v_z}_{\text{(i) prod.}} + \underbrace{\tau \epsilon^U v_U + \tau^D \epsilon v_D}_{\text{(ii) existing links to } U,\ D} + \underbrace{v_{\Delta U}(\tau) + v_{\Delta D}(\epsilon)}_{\text{(iii) option of future links to } U,\ D}, \tag{14}$$

which is achieved at equilibrium policies including a common innovation rate $\boldsymbol{\mu} > 0$ *and type-specific search and matching policies* $\boldsymbol{\lambda}(\tau)$ *and* $\boldsymbol{\beta}(\epsilon)$ *as strictly increasing functions in respective types,*

$$\boldsymbol{\mu} = (c'_\mu)^{-1}\left((1-b)\frac{v_z}{w}\right), \quad \boldsymbol{\lambda}(\tau) = (c'_\lambda)^{-1}\left(\tau \overline{\boldsymbol{\epsilon}_u}\frac{v_U}{w}\right), \quad \boldsymbol{\beta}(\epsilon) = (c'_\beta)^{-1}\left(\frac{\epsilon}{\overline{\boldsymbol{\beta}}}\overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}\frac{v_D}{w}\right), \tag{15}$$

where $\overline{\boldsymbol{\lambda}}$, $\overline{\boldsymbol{\beta}}$, $\overline{\boldsymbol{\tau}_d}$, *and* $\overline{\boldsymbol{\epsilon}_u}$ *are the sufficient statistics for* $\boldsymbol{M}$ *with the forms given in Eqs. (10) and (12). The common value coefficients* v_z, v_U, *and* v_D *satisfy*

$$v_z = \frac{\pi}{\rho + \boldsymbol{g}}, \qquad v_U = \frac{1}{\rho + \delta}\left(\boldsymbol{\mu}(1-b)v_z - w c_\mu(\boldsymbol{\mu})\right), \qquad v_D = \frac{1}{\rho + \delta}\boldsymbol{\mu} b v_z, \tag{16}$$

where the aggregate growth rate is given in Eq. (13), and type-specific $v_{\Delta U}(\cdot)$ *and* $v_{\Delta D}(\cdot)$ *are strictly increasing and* convex *functions, given as*

$$v_{\Delta U}(\tau) = \frac{1}{\rho}\left(\tau \boldsymbol{\lambda}(\tau)\overline{\boldsymbol{\epsilon}_u} v_U - w c_\lambda(\boldsymbol{\lambda}(\tau))\right), \quad v_{\Delta D}(\epsilon) = \frac{1}{\rho}\left(\frac{\epsilon \boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}}\overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}} v_D - w c_\beta(\boldsymbol{\beta}(\epsilon))\right). \tag{17}$$

Proposition 2 establishes that a firm's value nontrivially depends on its existing suppliers and users, as well as its potential to find new suppliers and users in the future; it also demonstrates the incentives for firms to innovate and to form new

links. Note that, in equilibrium, $v^{up}(\tau, \epsilon^U) = \tau\epsilon^U v_U + v_{\Delta U}(\tau)$ solves Eq. (8), and $v^{down}(\epsilon, \tau^D) = \tau^D \epsilon v_D + v_{\Delta D}(\epsilon)$ solves Eq. (9). This proposition offers several insights.

First, Eq. (14) shows that firm value comes from multiple sources, including (i) the product market, (ii) the existing links to upstream suppliers and to downstream users, and (iii) the option to form new links in the future. Second, Eqs. (15) to (17) illustrate that product-market profitability is the ultimate source of all the firm value components and drives the innovation and link-formation decisions. These decisions determine aggregate growth $\boldsymbol{g}$, which feeds back to the product-market profitability in equilibrium. Third, search frictions in forming new links lead to local granularity in decentralized markets for knowledge use, giving rise to local market powers in the allocation of innovation surplus; in turn, values of linked firms are also connected due to surplus sharing while the link lasts. Fourth, Eq. (15) demonstrates the search-attraction complementarity in the sense that firms search with greater intensities if they expect to find suppliers of better input types $\overline{\boldsymbol{\epsilon}_u}$, and firms exert higher attraction efforts if they expect better users $\overline{\boldsymbol{\tau}_d}$. Lastly, Eq. (17) also reveals that firm value not only strictly increases in own types τ and ϵ but are also strictly convex in them due to type-dependent search and attraction decisions. It means that network dynamics may amplify the type heterogeneity into more pronounced value heterogeneity.

4.1.3 What about within- and across-firm resource allocations?

Firms hire research labor to innovate and to form new links in the network. At any time t, a firm of type (τ, ϵ) and weighted supplier count ϵ^U hires a total flow of research labor $R(\tau, \epsilon, \epsilon^U) := \tau\epsilon^U c_\mu(\boldsymbol{\mu}) + c_\lambda(\boldsymbol{\lambda}(\tau)) + c_\beta(\boldsymbol{\beta}(\epsilon))$, increasing in τ, ϵ, and the endogenously evolving ϵ^U. Therefore, the within-firm allocation of research labor changes over time and depends on the set of suppliers it is matched to at the moment. When the firm has a larger and better set U of suitable inputs, it assigns a greater proportion of resources to the exploitation of U and a smaller proportion to finding new inputs. In the aggregate, the demand for research labor is $\boldsymbol{R} := \boldsymbol{X} c_\mu(\boldsymbol{\mu}) + \int c_\lambda(\boldsymbol{\lambda}(\tau)) dF_\tau + \int c_\beta(\boldsymbol{\beta}(\epsilon)) dF_\epsilon$, where the match-quality-weighted average degree $\boldsymbol{X}$ is introduced in Proposition 1. It implies that a denser knowledge network with better links requires more resource to build and more resource still to exploit.

4.2 Properties of the equilibrium: comparative statics

Propositions 1 and 2 suggest that endogenous link formation can serve as a natural amplification mechanism. That is, changes in model primitives that affect innovation incentives have impacts on growth $\boldsymbol{g}$ not only through the direct impact on innovation rate $\boldsymbol{\mu}$ but also the search and attraction policies. Meanwhile, link formation incentives have the equilibrium effect on $\boldsymbol{\mu}$ as well, indirectly affecting $\boldsymbol{g}$. Proposition 3 formalizes these discussions with comparative statics. For simpler exposition, we impose specific forms of cost functions. The logic shall apply more generally.

Proposition 3 (Comparative statics)**.** *Assume* $c_\mu(\mu) \propto \mu^{1+\chi_\mu}$, $c_\lambda(\lambda) \propto \lambda^{1+\chi_\lambda}$, *and* $c_\beta(\beta) \propto \beta^{1+\chi_\beta}$, *such that* $\chi_\mu, \chi_\lambda, \chi_\beta > 0$. *The unique stationary equilibrium features* $\overline{\boldsymbol{\tau}_d} = \frac{\int \tau^{1+1/\chi_\lambda} dF_\tau}{\int \tau^{1/\chi_\lambda} dF_\tau}$ *and* $\overline{\boldsymbol{\epsilon}_u} = \frac{\int \epsilon^{1+1/\chi_\beta} dF_\epsilon}{\int \epsilon^{1/\chi_\beta} dF_\epsilon}$ *irrespective of other primitives, such that* $\overline{\boldsymbol{\tau}_d}$ *and* $\overline{\boldsymbol{\epsilon}_u}$ *strictly decrease in* χ_λ *and* χ_β, *respectively. In the aggregate,*

1. *a higher return to innovation results in faster long-run growth* $\boldsymbol{g}$, *such that*

$$\frac{d\ln \boldsymbol{g}}{d\ln \pi} = -\frac{d\ln \boldsymbol{g}}{d\ln w} = \frac{d\ln \boldsymbol{g}}{d\ln(1-b)} = \frac{d\ln \boldsymbol{\mu}}{d\ln \pi} + \frac{d\ln \boldsymbol{x}}{d\ln \pi} = \frac{\frac{1}{\chi_\mu} + \frac{1}{\chi_\lambda}\frac{\chi_\mu+1}{\chi_\mu}}{1 + \left(\frac{1}{\chi_\mu} + \frac{1}{\chi_\lambda}\frac{\chi_\mu+1}{\chi_\mu}\right)\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}} > 0, \tag{18}$$

   $\forall \pi, w > 0$ *or* $b \in (0,1)$, *where network formation has an* amplifying *effect;*

2. *increased network's link fragility* δ *reduces long-run growth* $\boldsymbol{g}$, *such that*

$$\frac{d\ln \boldsymbol{g}}{d\ln \delta} = \frac{d\ln \boldsymbol{\mu}}{d\ln \delta} + \frac{d\ln \boldsymbol{x}}{d\ln \delta} = -\frac{1 + \frac{1}{\chi_\lambda}\frac{\delta}{\rho+\delta}}{1 + \left(\frac{1}{\chi_\mu} + \frac{1}{\chi_\lambda}\frac{\chi_\mu+1}{\chi_\mu}\right)\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}} < 0, \tag{19}$$

   $\forall \delta > 0$, *where network formation has a* dampening *effect when* δ *is small and an* amplifying *effect when* δ *is large.*

Proposition 3 shows how aggregate growth responds to the interacting innovation incentives and network formation incentives in equilibrium. The proof in Appendix A.3 explicitly compares with the exogenous-network case. As intuited before, Eq. (18) shows that endogenous network formation always *amplifies* the impact of changes in the return to innovation on growth $\boldsymbol{g}$. The reason is that the return to link formation always relies on and increases in the return to innovation. Note that a

greater $\boldsymbol{g}$ reduces the return to innovation, which is the usual business stealing effect in equilibrium, but it cannot undo the amplification effect.

A particular implication of Eq. (18) is the amplified negative impact on growth when the supplier's share b of innovation surplus rises. It suggests that policies such as IP protection may work in the opposite direction to stimulating growth. The reason is as follows. A larger supplier share b hurts the user's return to innovation and reduces $\boldsymbol{\mu}$. In turn, it hurts the innovator's incentives to find new suppliers. Meanwhile, a larger b is ineffective in driving up the quality of input in use $\overline{\boldsymbol{\epsilon}_u}$. Although firms as knowledge input suppliers have greater incentives now to attract new users, the extreme congestion effect renders it useless in producing more matches. Nonetheless, $b > 0$ is necessary for a nontrivial attraction policy, which ensures $\overline{\boldsymbol{\epsilon}_u} > \bar{\epsilon}$.

Eq. (19) shows that whether endogenous network formation *amplifies* or *dampens* the effect of changes in link destruction rate δ on growth $\boldsymbol{g}$ depends on the relative scale of δ. Smaller δ is associated with greater link formation incentives as firms expect to benefit from each user-supplier link for longer. Compared with an exogenous network dynamics setup, such incentives at a small δ alleviate the negative impact of increases in δ on $\boldsymbol{g}$. As δ gets larger, firms have fewer incentives to form new links, reducing $\boldsymbol{g}$ further.

4.3 Equilibrium degree evolution and distributions

Solving for the equilibrium needs only the aggregated information of the network, i.e., the unweighted average degree $\boldsymbol{x}$ and the weighted $\boldsymbol{X}$, both of which are constant and direction-irrelevant. However, at the individual level, the network always evolves with new links born and old ones gone, resulting in within-type firm heterogeneity. Furthermore, each firm's U and D are generically asymmetric.

Proposition 4 (Type-specific degree evolution and distribution)**.** *At a stationary equilibrium with average degree $\boldsymbol{x}$, consider a firm of type (τ, ϵ):*

1. *conditional on the output type τ, its supplier count (outdegree) $x^U = |U|$ evolves as a birth-death process with birth rate $\boldsymbol{\lambda}(\tau)$ and death rate δx^U, and the ergodic distribution is $x^U \mid \tau \sim \text{Poisson}\left(\boldsymbol{x}\boldsymbol{\lambda}(\tau)/\overline{\boldsymbol{\lambda}}\right)$;*

2. *conditional on the input type ϵ, its user count (indegree) $x^D = |D|$ evolves as a*

birth-death process with birth rate $\overline{\boldsymbol{\lambda}}\boldsymbol{\beta}(\epsilon)/\overline{\boldsymbol{\beta}}$ *and death rate* δx^D*, and the ergodic distribution is* $x^D \mid \epsilon \sim \text{Poisson}\left(\boldsymbol{x}\boldsymbol{\beta}(\epsilon)/\overline{\boldsymbol{\beta}}\right)$.

Proposition 4 shows the within- and across-type heterogeneity in firms' supplier counts and user counts. The Poisson distributions result from random matching and random link destruction. On average, firms searching at higher intensities $\boldsymbol{\lambda}(\tau)$ have proportionally more suppliers and exhibit larger dispersion in supplier counts, and firms exerting higher attraction efforts $\boldsymbol{\beta}(\epsilon)$ have proportionally more users and show larger dispersion in user counts. A firm with many suppliers may not have many users even in expectations, unless τ and ϵ are positively correlated across firms.

Compared with unweighted degrees x^U and x^D, the match-quality-weighted supplier count $\tau\epsilon^U$ and user count $\tau^D\epsilon$ are more relevant for firm value, such that $\tau\epsilon^U$ governs the evolution of relative productivity z, and $\tau^D\epsilon$ determines the rent flow extracted from users. The evolution of $\epsilon^U \mid \tau$ and $\tau^D \mid \epsilon$ follows that of $x^U \mid \tau$ and $x^D \mid \epsilon$ but involves additional randomness. Firm productivity jumps at rate $\boldsymbol{\mu}$ of successful innovation with the step size depending on the firm's own τ and its suppliers' ϵ^U at the moment. Therefore, output type τ governs the long-run average of each firm's productivity relative to the aggregate growth trend. Corollary 1 helps to tie the heterogeneity in firm types to that in measures of firm size, such as relative productivity z, profit from sales πz, innovation output $\boldsymbol{\mu}\tau\epsilon^U$, and firm value $V(\tau, \epsilon, z, \epsilon^U, \tau^D)$.

Corollary 1 (Firm type and size)**.** *At a stationary equilibrium and in the long run, a firm's average match-quality-weighted supplier count is* $\mathbb{E}[\tau\epsilon^U \mid \tau] = \overline{\boldsymbol{\epsilon}_u}\boldsymbol{x}\tau\boldsymbol{\lambda}(\tau)/\overline{\boldsymbol{\lambda}}$*; its average match-quality-weighted user count is* $\mathbb{E}[\tau^D\epsilon \mid \epsilon] = \overline{\boldsymbol{\tau}_d}\boldsymbol{x}\epsilon\boldsymbol{\beta}(\epsilon)/\overline{\boldsymbol{\beta}}$*; its average relative productivity* z *is* $\mathbb{E}[z \mid \tau] = \mathbb{E}[\tau\epsilon^U \mid \tau]/\boldsymbol{X} = \tau\boldsymbol{\lambda}(\tau)/(\overline{\boldsymbol{\tau}_d} \cdot \overline{\boldsymbol{\lambda}})$.

4.4 Taking stock: empirical findings revisited

The benchmark model captures the majority of the findings in Section 2. With one firm representing each industry, a new pathway corresponds to a new link in the model. We pinpoint each finding's counterpart in the model. First, pathways are formed infrequently. In the model, new links emerge at the aggregate rate $\overline{\boldsymbol{\lambda}}$ due to search frictions. Second, pathway citations tend to connect large firms and high-quality innovations (Facts 1 and 2). In the model, the likelihood that a particular searching firm of type τ forms a link with a particular attracting firm of type ϵ' is

$\boldsymbol{\lambda}(\tau)\boldsymbol{\beta}(\epsilon')/(\overline{\boldsymbol{\lambda}\boldsymbol{\beta}})$, which strictly increases in τ and ϵ'. Firm types govern firm sizes. Third, firm innovation and real performance improve after forming pathways (Facts 3 and 4). In the model, when a new link is formed between a user of type τ (pathfinder) and a supplier of type ϵ' (receiver), the finder's value increases by $\Delta V_{\text{finder}} = \tau\epsilon' v_U$, and the receiver's value increases by $\Delta V_{\text{receiver}} = \tau\epsilon' v_D$. The benchmark model does not have firm heterogeneity within an industry, so pathways' spillover effects on either end's peer firms are missing. The extended model in Section 5 takes care of it.

4.5 Discussion of model assumptions

Several assumptions deserve more discussions, all of which help simplify the analysis.

Productivity jumps. The crucial assumption is that, conditional on the arrival of a successful innovation, a firm's productivity jump size $\sum_{u\in U}\tau\epsilon_u \boldsymbol{Z}(t)$ depends on the available suppliers' types $\{\epsilon_u\}_{u\in U}$ and the aggregate productivity $\boldsymbol{Z}(t)$ but not the suppliers' productivity levels $\{Z_u(t)\}_{u\in U}$. This is plausible because the productivity level $Z_u(t)$ of supplier u in producing physical goods or services at the time should not have direct impact on the knowledge input users' innovation outcome, but the persistent knowledge applicability ϵ_u does. Moreover, this assumption is critical for model tractability because it renders a firm's own or its neighbors' endogenous and network-dependent productivity levels irrelevant for link formation incentives. Otherwise, one would need to track a firm's value changes in response to the endogenous productivity changes of indefinitely many firms. See Appendix A.2.1 for more details. Modeling endogenous network dynamics is typically challenging in part because of these higher-order impacts on agent payoffs. Hence, reducing the "depth" of the local network structure is usually necessary for tractability. One option with unknown performance is to deviate from rationality and force agents to be oblivious to state changes more than some predetermined $k \in \mathbb{N}$ steps away. In contrast, the current setup not only ensures analytical convenience but preserves full rationality. It also preserves the relevance of both the local network (discrete U and D) and the aggregate network (summarized by $\boldsymbol{X}$) to a firm's payoff.

Separable suppliers. The assumption of each supplier's separable contribution to a firm's payoff has two desirable consequences. One is the separable functional form of firm value in Proposition 2. It is transparent how the product-market profitability

determines innovation incentives and payoffs, which, in turn, govern the search and attraction incentives to form new links in the knowledge network. It is also clear how the aggregate network's density and quality feed back to the equilibrium profitability via $\boldsymbol{X}$ in $\boldsymbol{g}$ as shown in Proposition 1. The other consequence regards the allocation of innovation surplus among an innovating firm and its suppliers. With separable contribution of each supplier, the surplus sharing outcome is the least sensitive to multilateral surplus allocation protocols. We caution that the additive separability in the productivity jump should not be interpreted as perfect substitution of knowledge inputs across industries. The fact that firms must search for each applicable input reflects a lack of substitutability. Nonetheless, we believe that the interplay of non-separable suppliers, surplus allocation protocol, and the network dynamics and structure is fascinating and worthy of future research. Indeed, Appendix A.2.2 shows that it is feasible to incorporate supplier interaction in this model at some cost of analytical tractability. However, it is beside the primary goal of the present paper.

Rent extraction from downstream users. The assumption of rent extraction by suppliers at each successful downstream innovation intends to capture the empirical observation that firms benefit from being on the receiving ends of pathways. Rent extraction is a modeling device, and it is consistent with the direct transfers among firms associated with knowledge use in reality, such as royalty payments and licensing fees. It is also in line with the scenario that uses of knowledge inputs lead to uses of goods or services as production inputs. In either case, supplier revenue rises when user profitability rises due to successful innovations. Whether direct transfers are involved is less crucial. Indeed, an equivalent and more reduced-form specification is to assume simultaneous jumps in the innovator's and its existing suppliers' relative productivities upon each successful innovation, such that the innovator's z jumps to $z + (1-b)\tau\epsilon^U$, and each supplier u's z_u jumps to $z_u + b\tau\epsilon_u$, $\forall u \in U$. The equilibrium analysis is essentially unchanged.[10]

[10] Of course, the evolution of individual productivity will be adjusted accordingly. Consequently, the conditional expectation of relative productivity in Corollary 1 becomes $\mathbb{E}[z \mid \tau, \epsilon] = (1-b)\frac{\tau\boldsymbol{\lambda}(\tau)}{\overline{\boldsymbol{\tau}_d}\cdot\overline{\boldsymbol{\lambda}}} + b\frac{\epsilon\boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\epsilon}_u}\cdot\overline{\boldsymbol{\beta}}}$. However, the rest of the equilibrium outcome remains unaffected.

5 Extended Model: Knowledge Path Externality

The setup of the extended model differs from the benchmark in two major aspects. First and foremost, each industry now has multiple firms, which is necessary to capture the spillover effects of pathway formation. In particular, firms free ride on others' efforts to form new links in the *industry-level* network, referred to as the *knowledge path externality*. Second, researcher wage rate w and firms' flow profits πz are endogenously determined in a general equilibrium with household optimization and firm production and pricing decisions. As we highlight in Proposition 5, aggregate growth $\boldsymbol{g}$ in equilibrium is still tightly linked to the endogenous knowledge network, but the *knowledge path externality* affects the network's *density* and *quality*.

The extended model is built to follow the benchmark as closely as possible. The equilibrium includes a common innovation rate $\boldsymbol{\mu} > 0$ and type-specific search and attraction policies $\boldsymbol{\lambda}(\tau)$ and $\boldsymbol{\beta}(\epsilon)$ as in the benchmark case, although more parameters are introduced for the quantitative exercises in Section 6. To be brief, we only specify the changes made to the setup and discuss their impacts on the equilibrium outcome of interest. Omitted details are in Appendix B.

5.1 The modified setup to close the model

Each industry out of the unit measure of them $j \in [0, 1]$ has a fixed and finite number $n_j \geq 2$ of risk-neutral firms, $n_j \in \mathbb{N}$. The cross-sectional distribution of industry sizes has a cdf F_n with mean N. A firm is indexed by (j, f) or simply a generic f for brevity, and $f \in j$ is shorthand for a firm f in industry j. Firms in j share a common input type $\epsilon_j \sim F_\epsilon$ but differ in output types $\tau_{jf} \sim F_\tau$, $f \in j$. For simplicity, assume that an industry's size and its firm types are independently distributed. Each firm produces a distinct product. The total measure of firms or products is fixed at N. Households consume, work, and own all firms. An infinitely lived representative household summarizes their behavior. Product markets are monopolistically competitive. Labor markets are perfectly competitive.

Firms in a given industry j share *similar* knowledge, in the sense that (i) they are (imperfect) substitutes as suppliers of knowledge input j of type ϵ_j, and (ii) they can innovate using the same knowledge inputs at respective output types τ_{jf}, $f \in j$.

In other words, all firms in a given industry j share a common upstream set $U_j(t)$ of supplier industries and a common downstream set $D_j(t)$ of user industries at any time t. Endogenous network dynamics mean the *industry-level* evolution of U and D for each industry as a result of individual *firm-level* decisions.

Households. A representative household has flow utility $\ln C(t) - \varsigma \frac{R(t)^{1+1/\sigma}}{1+1/\sigma}$. The consumption compound is $C(t) = \left(\int_0^1 \sum_{f \in j} c_{jf}(t)^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}$, with a common elasticity of substitution $\nu > 1$ across all products of mass F. $R(t)$ is the supply of research labor, with $\varsigma > 0$ and Frisch elasticity $\sigma \geq 0$. Production labor is in fixed supply L and serves as the *numeraire*. Discount rate is $\rho > 0$. Given product prices $p_{jf}(t)$, researcher wage rate $w(t)$, and total dividend payment $\Pi(t)$ at t, household budget constraint reduces to $P(t)C(t) = \int \sum_{f \in j} p_{jf}(t) c_{jf}(t) dj = L + w(t)R(t) + \Pi(t)$, where $P(t)$ is the ideal price index.

Firm production and product pricing. Firms hire production labor at unit wage rate to produce at each instant t. The production function is $\widehat{O}_{jf}(t) = Z_{jf}(t)^{\frac{1}{\nu-1}} \widehat{l}_{jf}(t)$, where $\widehat{O}_{jf}(t)$ is firm (j,f)'s output flow, $\widehat{l}_{jf}(t)$ is the production labor hired, and $Z_{jf}(t)^{\frac{1}{\nu-1}}$ is its rescaled productivity. $Z_{jf}(t)$ grows due to innovation. The firm sets the price $\widehat{p}_{jf}(t)$ under monopolistic competition.

Household optimization and product-market outcome. We set up the household problem and firm production and pricing problem as such to endogenize the instantaneous profit πz and the researcher wage rate w assumed in the benchmark model. Choosing production labor in fixed supply as the numeraire implies that the risk-free interest rate is simply ρ. The supply of research labor is $R(\mathbf{w}) := (\mathbf{w}/\varsigma)^\sigma$. Firms optimally price their products at a constant markup over the marginal cost. When firm (j,f)'s productivity level at t is $Z_{jf}(t) = Z$, its optimal price is $p^*_{jf}(t) = \frac{\nu}{\nu-1} Z^{-\frac{1}{\nu-1}}$, and its flow profit is $\pi^*_{jf}(t) = \boldsymbol{\pi} z$, where $\boldsymbol{\pi} = \frac{1}{N} \frac{L}{\nu-1}$ is the average profit per firm, and $z := Z/\boldsymbol{Z}(t)$ is firm (j,f)'s productivity *relative* to the cross-sectional firm-level mean $\boldsymbol{Z}(t)$ at t, where $\boldsymbol{Z}(t) := \frac{1}{N} \int_0^1 \sum_{f' \in j'} Z_{j'f'}(t) dj'$ with some abuse of notation. Aggregation yields $\boldsymbol{C}(t) = (N \boldsymbol{Z}(t))^{\frac{1}{\nu-1}} L$, and the growth rate of aggregate consumption is $\frac{\boldsymbol{g}}{\nu-1}$ when the average productivity grows at rate $\boldsymbol{g} := \dot{\boldsymbol{Z}}(t)/\boldsymbol{Z}(t)$.

5.2 Firm-level decisions and an industry-level network

The rest of the firm problem pertaining to the knowledge network is similar to the benchmark model but tweaked to accommodate within-industry firm heterogeneity and interactions. Consider an industry j of $n_j = n$ firms with a common input type $\epsilon_j = \epsilon$ and individual output types $(\tau_f)_{f\in j}$. At time t, j has a set $U_j(t) = U$ of upstream industries with a combined input type $\epsilon^U := \sum_{u\in U} \epsilon_u$, where each ϵ_u is shared by all the n_u firms in industry u, $\forall u \in U$. Meanwhile, j has a set $D_j(t) = D$ of downstream industries such that industry d has n_d firms, $\forall d \in D$. To describe firm decisions, we focus on a particular firm $f \in j$ of output type $\tau_f = \tau$ and relative productivity z at time t. The firm f takes as given, among other things, the growth rate $\boldsymbol{g}$, the profit coefficient $\boldsymbol{\pi}$, and the researcher wage rate $\mathbf{w}$ in equilibrium. We build toward a separable firm value function $V(\cdot)$ with the term zv_z satisfying Eq. (2).

Innovation and surplus sharing. To account for growth *without* network origins, we assume that firms can always use own-industry knowledge as an innovation input. Conditional on a successful innovation at a chosen $\widehat{\mu}$, the innovating firm's relative productivity jumps from z to $z + \tau\big(\epsilon^U + \eta\epsilon\big)$, where the parameter $\eta \geq 0$ is a potential "home bias" towards using own-industry inputs. Larger η attributes more growth to non-network origins. Achieving $\widehat{\mu}$ requires the firm to hire $\tau\big(\epsilon^U + \eta\epsilon\big)c_\mu(\widehat{\mu})$ units of research labor at wage rate $\mathbf{w}$.

The specific supplier firm of each upstream industry $u \in U$ and own industry j is randomly drawn and matched to firm f every time f innovates. That is, conditional on f's successful innovation, any firm in $u \in U$ becomes f's matched supplier of input u with probability $1/n_u$ and match quality $\tau\epsilon_u$, and any other firm in j becomes f's matched supplier of input j with probability $1/(n-1)$ and match quality $\tau\eta\epsilon$. If firm f were to switch to an unmatched firm as the supplier of input $u \in U$ or j, it would incur a switching cost in match quality $\kappa\tau\epsilon_u$ or $\kappa\tau\eta\epsilon$, where the cost proportion is $\kappa \in (0, 1]$. With this understanding, the innovating firm f shares the gross innovation surplus $S(\tau, \epsilon^U + \eta\epsilon) = \tau\big(\epsilon^U + \eta\epsilon\big)v_z$ with the *matched* suppliers according to the Shapley value solution. Parameter κ captures the extent to which knowledge within the same industry is *substitutable*, and it reflects the supplier's power in the local knowledge-use market. A smaller κ means more substitutable knowledge inputs within an industry and lower market power of suppliers, raising the

innovator's surplus while reducing the suppliers'. It is straightforward to show that the matched supplier firm of input $u \in U$ gets a share $b = \kappa/2$ of pairwise surplus $S(\tau, \epsilon_u)$, the matched supplier firm in own industry j gets the same share $b = \kappa/2$ of $S(\tau, \eta\epsilon)$, and the innovating firm f gets the remaining $(1-b)S(\tau, \epsilon^U + \eta\epsilon)$. See Appendix B.3 for specifics. It follows that the equilibrium innovation rate $\boldsymbol{\mu} > 0$ still satisfies the expression in Eq. (15) evaluated at $w = \mathbf{w}$. In equilibrium, firm f becomes the matched supplier to any downstream innovator with probability $1/n$ and expects to extract from downstream industries a flow rent $\boldsymbol{\mu}\frac{1}{n}\sum_{d\in D}\sum_{f'\in d} bS(\tau_{df'}, \epsilon)$. It becomes the the matched supplier to another firm in own industry with probability $1/(n-1)$, so the expected rent from own industry is $\boldsymbol{\mu}\frac{1}{n-1}\sum_{f'\in j\setminus\{f\}} bS(\tau_{f'}, \eta\epsilon)$.

Forming and sharing new links. The knowledge network evolves at the industry level, such that all firms in j gain or lose members of U and D at the same time. Random link destruction is exogenous at Poisson rate $\delta > 0$.

Each firm chooses an individual search rate $\widehat{\lambda} \geq 0$ and an individual attraction intensity $\widehat{\beta} \geq 0$ at the respective costs in researcher wage $\mathbf{w}c_\lambda(\widehat{\lambda})$ and $\mathbf{w}c_\beta(\widehat{\beta})/n$. The adjustment $1/n$ in the attraction cost means that it is cheaper for firms in a larger industry to attract users. It mirrors each firm's $1/n$ chance of becoming a matched supplier to a downstream firm, ensuring a scale-free optimal attraction policy. Let the individual search and attraction policies in equilibrium be $\boldsymbol{\lambda} : [\tau_0, \tau_1] \to \mathbb{R}_+$ and $\boldsymbol{\beta} : [\epsilon_0, \epsilon_1] \to \mathbb{R}_+$, cross-sectionally aggregated at $N\overline{\boldsymbol{\lambda}}$ and $N\overline{\boldsymbol{\beta}}$. In equilibrium, industry j gets a new upstream industry at rate $\sum_{f'\in j}\boldsymbol{\lambda}(\tau_{f'})$ and a new downstream industry at rate $N\overline{\boldsymbol{\lambda}}n\boldsymbol{\beta}(\epsilon)/(N\overline{\boldsymbol{\beta}})$.

Every link connecting two industries is an industry-wide public good in either industry. Whenever one firm $f \in j$ forms a new link in either direction, U or D expands, and all the other firms $f' \neq f \in j$ immediately benefit from it as well. However, firm f does not take into account the expected value gain by the other $n-1$ firms in j when choosing its search rate $\widehat{\lambda}$ or attraction level $\widehat{\beta}$.

Formation of each industry-to-industry link (pathway) always directly involves a pair of firms, i.e., the searching firm as the pathfinder and the attracting firm as the path-receiver. To match the data, we allow for a potential "path premium" as the extra reward exclusive to this pair, captured by an additional parameter $\gamma \geq 0$. Specifically, the pathfinder (j, f)'s relative productivity jumps immediately from z_{jf}

to $z_{jf} + \gamma\tau_{jf}\epsilon_i/n_i$ when finding the receiver firm (i, h) as the input supplier, $\forall h \in i$. Note that such productivity jumps directly contribute to aggregate growth $\boldsymbol{g}$. The jump size is again adjusted by the upstream industry's size $1/n_i$ to remove the scale effect. The pathfinder and receiver firms split the pair-specific surplus $S(\tau, \gamma\epsilon_i/n_i) = \tau\gamma\epsilon_i/n_i v_z$ by *symmetric Nash* bargaining, knowing that (j, f) has the option to switch to another firm $h' \in i$ at the proportional cost of κ. It follows that the receiver firm in i, i.e., the supplier, still gets the share $b = \kappa/2$ of $S(\tau, \gamma\epsilon_i/n_i)$, and the finder f, i.e., the user, gets the remaining $1 - b = 1 - \kappa/2$.

5.3 Equilibrium characterizations and implications

We focus on the equilibrium outcome that differs from the benchmark model and is relevant for the quantitative exercise. The primary variable of interest is aggregate growth, now driven by multiple forces with and without network origins. Therefore, $\boldsymbol{g}$ has a richer expression more suitable for the quantitative exercise.

Proposition 5. *The aggregate productivity growth rate $\boldsymbol{g}$ in equilibrium is given as*

$$\boldsymbol{g} = \underbrace{\boldsymbol{\mu}\eta\overline{\tau\epsilon}}_{\text{(i) without link}} + \underbrace{\boldsymbol{\mu}\boldsymbol{X}}_{\text{(ii) existing links}} + \underbrace{\overline{\boldsymbol{\tau}_d}\frac{\gamma\overline{\boldsymbol{\lambda}}}{N}\overline{\boldsymbol{\epsilon}_u}}_{\text{(iii) path premium}}, \tag{20}$$

where the match-quality-weighted average degree $\boldsymbol{X}$ in (ii) is given as

$$\boldsymbol{X} = \Big(\underbrace{\tfrac{1}{N}\overline{\boldsymbol{\tau}_d}}_{\text{pathfinder}} + \underbrace{\tfrac{N-1}{N}\overline{\tau}}_{\text{peer}} \Big)\boldsymbol{x}\overline{\boldsymbol{\epsilon}_u} \in (\overline{\tau}\boldsymbol{x}\overline{\epsilon}, \overline{\boldsymbol{\tau}_d}\boldsymbol{x}\overline{\boldsymbol{\epsilon}_u}) \tag{21}$$

such that the unweighted average degree is $\boldsymbol{x} = N\overline{\boldsymbol{\lambda}}/\delta$.

In Eq. (20), the essential network-driven component (ii) of growth is inherited from Eq. (13) of the benchmark model, but the match-quality-weighted average degree $\boldsymbol{X}$ given in Eq. (21) is subject to the *knowledge path externality*. As before, the expected output type of a pathfinder firm is $\overline{\boldsymbol{\tau}_d} > \overline{\tau}$, and the expected input type of a path-receiver firm (and its peers with the same type) is $\overline{\boldsymbol{\epsilon}_u} > \overline{\epsilon}$. However, the expected output type of a peer firm to any pathfinder is $\overline{\tau}$. While only one firm can be the finder of a given link, all of its peers in the same industry can use the link; the average number of peer firms in each industry is $N - 1$. In expectations, the average

output type of the user industry becomes $\frac{1}{N}\overline{\boldsymbol{\tau}_d}+\frac{N-1}{N}\overline{\tau}$, which is below the pathfinder's average $\overline{\boldsymbol{\tau}_d}$. Network density $\boldsymbol{x}$ appears to be a simple rescale by the measure of firms. However, individual firms' search policies comprising $\overline{\boldsymbol{\lambda}}$ are also under the impact of the *knowledge path externality*, discussed below. Therefore, both the *density* and *quality* of the network are affected.

Component (i) of Eq. (20) is the innovation outcome attributed to the use of within-industry knowledge input, which is considered as growth without network origins. Component (iii) reflects the extra productivity jump at the formation of new links due to the "path premium". Qualitatively, terms (i) and (iii) have little impact on the model's structure, and they disappear at $\eta = 0$ and $\gamma = 0$. Quantitatively, however, they determine the extent to which aggregate growth can be attributed to endogenous knowledge network dynamics.

We turn to individual firm decisions. Here, we again highlight the *knowledge path externality* that arises from link-sharing within an industry. When a firm of output type τ searches at intensity $\widehat{\lambda}$, it is expected to generate an industry-wide flow value of $\widehat{\lambda}\overline{\boldsymbol{\epsilon}_u}\Big(\tau\frac{\gamma}{N}\left(1-\frac{\kappa}{2}\right)v_z+\tau v_U+\sum_{\tau'\in\tau_{-f}}\tau' v_U\Big)$, where τ_{-f} contains the output types of the $n-1$ other firms in the same industry, and v_U is the marginal value of match-quality-weighted supplier counts. The value gain $\widehat{\lambda}\overline{\boldsymbol{\epsilon}_u}\sum_{\tau'\in\tau_{-f}}\tau' v_U$ goes to the other $n-1$ firms and is not internalized in the choice of $\widehat{\lambda}$. Meanwhile, firms share the downstream links to extract innovation surplus from users, and they ignore the spillovers to other firms when choosing attraction levels. This positive externality partially offsets the negative congestion effect when firms compete for user attention.

Two other consequences of the *knowledge path externality* are worth mentioning. One regards the value of an individual firm. The firm value function can be solved analogously to the benchmark with details relegated to Appendix B.2. Not only is the firm's value strictly increasing in any given peer's output type $\tau' \in \tau_{-f}$ such that $\partial V/\partial\tau' > 0$, but it is also super-modular in own τ and any peer's τ' such that $\partial^2 V/\partial\tau\partial\tau' > 0$. Relatedly, the long-run average of firm's relative productivity $\mathbb{E}[z \mid \tau, \tau_{-f}, \epsilon]$ is also strictly increasing and super-modular in τ and τ'. The intuition is that firms of better output types can better take advantage of new upstream links formed by others in innovation, so their values and long-run relative productivity levels are increasingly higher if their peers form more of these links. The other consequence

concerns within-industry resource allocation. A firm's demand for research labor flow totals at $R(\tau, \epsilon, n, \epsilon^U) = \tau(\epsilon^U + \eta\epsilon)c_\mu(\boldsymbol{\mu}) + c_\lambda(\boldsymbol{\lambda}(\tau)) + c_\beta(\boldsymbol{\beta}(\epsilon))/n$. Firms in the same industry share a common ϵ^U, so those of lower τ's are expected to allocate greater proportions of research labor toward exploiting existing upstream links rather than searching for new suppliers.

6 Quantitative Illustrations

To what extent do endogenous knowledge dynamics drive aggregate growth?[11] This section attempts to answer this question numerically after parameterizing the extended model. Appendix C contains omitted details.

Additional data. We obtain US R&D employment from OECD Statistics, total employment and corporate profit over GDP from Federal Reserve Economic Data, R&D investment over GDP ratio from National Center for Science and Engineering Statistics of National Science Foundation, R&D worker compensation over total R&D investment from National Center for Science and Engineering Statistics and Census Bureau, Business Enterprise Research and Development Survey. Other patent citation and granted application related moments come from USPTO's www.patentsview.org. We divide the data into two subsamples with equal time length, I. 1981–1997 and II. 1998–2014.

Identification and calibration. Parameter values are in Table 3. Unit time length is a year. Time discount is set at $\rho = 0.05$. Frisch elasticity of research labor supply is $\sigma = 0.82$ following Chetty et al. (2011). The fixed production labor supply is normalized at $L = 100\frac{\nu-1}{\nu}$, so the aggregate sales are $\boldsymbol{P}(t)\boldsymbol{C}(t) = 100$ in equilibrium. Assume exponential type distributions F_τ and F_ϵ, imposing $\overline{\tau} = \overline{\epsilon} > 0$ as they cannot be separately identified. Assume an innovation cost function $c_\mu(\mu) = \frac{c_\mu}{2}\mu^2$ with $c_\mu > 0$, where the quadratic shape follows Acemoglu et al. (2018). Assume a search cost function $c_\lambda(\lambda) = \frac{c_\lambda^{\chi_\lambda}}{1+\chi_\lambda}\lambda^{1+\chi_\lambda}$ with $c_\lambda, \chi_\lambda > 0$. Assume an attraction cost function $c_\beta(\beta) = \frac{1}{1+\chi_\beta}\beta^{1+\chi_\beta}$ with $\chi_\beta > 0$, where the scale parameter is normalized to be 1 as it cannot be identified without directly observing $\boldsymbol{\beta}(\epsilon)$ in the data.

Parameters (N, η, χ_β) have immediate empirical counterparts. The average number of firms N per industry is read directly from data. The "home-bias" η of input

[11] We thank Diego Restuccia for inspiring this exercise.

Table 3: Parameter values

	predetermined		externally calibrated			internally calibrated								
	ρ	σ	N	η	χ_β	ν	ς	κ	δ	c_μ	c_λ	χ_λ	$\bar{\epsilon}$	γ
I.	0.05	0.82	133	170	0.42	14.8	3.90	0.74	0.10	0.20	21.4	110	0.022	13743
II.	0.05	0.82	185	327	0.53	9.38	3.80	0.51	0.03	0.26	48.7	66.8	0.013	55591

I and II indicate the subsample, I. 1981–1997, II. 1998–2014. See main text for parameter descriptions.

use in innovation is set to be the average ratio of citations into own-industry (j to j) over average citations into any other cited industry (j to $j' \neq j$). To back out the shape parameter χ_β for the attraction cost, we leverage the following result.

Corollary 2. *Assume $c_\beta(\beta) \propto \beta^{1+\chi_\beta}$ and an exponential distribution for the input type ϵ. Then, $\overline{\boldsymbol{\epsilon}_u}/\bar{\epsilon} = 1 + 1/\chi_\beta$ in equilibrium.*

Therefore, we have $\chi_\beta = \left(\frac{\overline{\boldsymbol{\epsilon}_u}}{\bar{\epsilon}} - 1\right)^{-1}$. We do not directly observe the average type of inputs in use $\overline{\boldsymbol{\epsilon}_u}$ or the average input type $\bar{\epsilon}$. However, input type determines how often a piece of knowledge is used as an input in the model. In the data, citation flows approximate knowledge use. Therefore, we approximate the empirical counterpart to $\frac{\overline{\boldsymbol{\epsilon}_u}}{\bar{\epsilon}}$ as the ratio of average number of citations received per cited input patent (second-order indegree) over the number of citations per patent (first-order indegree).

The parameters to be calibrated internally are $\theta = (\nu, \varsigma, \kappa, c_\mu, \delta, \chi_\lambda, \bar{\epsilon}, c_\lambda, \gamma)$. We show exact identification step by step in Appendix C using the moments listed in Table 4, leveraging the closed-form equilibrium conditions. The intuition is as follows. The aggregate profit share identifies the household's elasticity of substitution ν among products. The researcher employment share identifies the disutility scale parameter ς. Researcher compensation as a fraction of total R&D cost identifies the parameter κ for supplier switching costs. Innovation frequency identifies the scale parameter c_μ in innovation cost. The average degree of industry-level knowledge network identifies the link destruction rate δ given the average probability of finding a pathway. The remaining four moments can be written as a system of four equations with the last four parameters as the unknowns: search cost scale c_λ and shape χ_λ, average type $\bar{\epsilon}$, and the path premium γ. Parameter values are picked by the generalized method of moments, such that θ solves $\min_\theta \sum_{i=1}^{9} \frac{2[m^{data}(i) - m^{model}(\theta,i)]^2}{m^{data}(i)^2 + m^{model}(\theta,i)^2}$.

Table 4: Goodness of fit: targeted moments

		I. 1981–1997		II. 1998–2014	
	description	data	model	data	model
(1)	R&D expenditure to GDP ratio (%)	2.49	3.04	2.62	3.65
(2)	researcher employment share (%)	0.66	0.67	0.86	0.86
(3)	profits to GDP ratio (%)	5.40	5.39	8.55	8.53
(4)	innovation rate (%)	9.59	9.59	6.18	6.16
(5)	firms' average innovation frequency (%)	50.4	50.5	53.5	53.4
(6)	researcher compensation share of total R&D (%)	43.4	45.7	53.6	58.5
(7)	firms' average probability of finding a pathway (%)	4.48	4.48	1.90	1.92
(8)	avg. degree of industry-level knowledge network	59.5	58.6	120	120
(9)	avg. pathfinder and receiver's sales increase (%)	1.50	1.51	1.50	1.50

Identification is shown in Appendix C. Computation of the model is straightforward with a simple algorithm sketched also in Appendix C.

Growth decomposition. Eq. (20) in Proposition 5 allows us to decompose the aggregate growth $\boldsymbol{g}$ measured as innovation rate according to the origins for each subsample. In turn, we can use it to analyze the changes in growth over the last few decades. Table 5 shows the results. Cross-category knowledge network evolution (ii)+(iii) contributes to more than one half of aggregate growth in each subsample. The vast majority of growth of network origins is associated with free-riding on others' search efforts, i.e., knowledge-path externalities. It is not so surprising given the number of firms innovating in each category. Productivity jumps at the relatively infrequent formation of new links also contributes a significant portion. Overtime, growth of any origin declines, with network evolution accounting for almost two thirds of the drop. The decline at new link formation is proportionally more severe.

Another exercise is to isolate the impact of each key parameter on growth and other equilibrium variables. Treating calibration I as the baseline, we set one parameter at a time to its value at II and report the change in equilibrium outcome in Table 6. The first row examines the decline in growth from an alternative angle. The other rows concern equilibrium variables closely related to growth. The drop in average efficiency type ($\overline{\tau} = \overline{\epsilon}$) is the main force of growth slowdown, although it elevates innovation frequency ($\boldsymbol{\mu}$) due to general equilibrium effect through researcher wage. Increases in innovation and search cost coefficients (c_μ and c_λ) damage growth as

Table 5: Growth decomposition based on Eq. (20) according to the growth origins

		(i) $\boldsymbol{\mu}\eta\overline{\tau\epsilon}$	(ii) $\boldsymbol{\mu X}$		(iii) $\overline{\boldsymbol{\tau}_d}\frac{\gamma\overline{\boldsymbol{\lambda}}}{N}\overline{\boldsymbol{\epsilon}_u}$
subsample	$\boldsymbol{g}$	no network	own search	others' search	new links
I. 1981–1997 (pp)	9.59	4.09	0.04	4.72	0.75
(% of $\boldsymbol{g}$)	100	42.6	0.38	49.2	7.84
II. 1998–2014 (pp)	6.16	2.86	0.02	3.01	0.28
(% of $\boldsymbol{g}$)	100	46.5	0.27	48.8	4.47
change, II−I (pp)	−3.43	−1.22	−0.02	−1.71	−0.48
(% of $\Delta\boldsymbol{g}$)	100	35.6	0.57	49.9	13.9

Table 6: Impact of each parameter on the equilibrium outcome (baseline = calibration I)

equilibrium	baseline	gross change	change in eqm. variable if one parameter takes its value at II							
variable	value at I	from I to II	ν	$\bar{\epsilon}$	κ	δ	c_μ	c_λ	χ_λ	χ_β
$\boldsymbol{g}$ (pp)	9.58	−3.62	+1.07	−4.37	+0.45	+5.14	−1.25	−1.87	−0.08	−0.46
$\boldsymbol{\mu}$ (pp)	50.5	+2.94	+6.08	+32.1	+2.57	−15.7	−7.14	+9.88	+0.34	+2.25
$\overline{\boldsymbol{\lambda}}$ (pp)	4.57	−2.65	+0.00	−0.02	+0.00	−0.02	+0.00	−2.51	−0.12	−0.00
$\boldsymbol{X}$ ($\times 10^2$)	9.80	−4.15	+0.01	−6.19	+0.00	+22.7	+0.01	−5.27	−0.25	−1.41

well. The lower risk of link destruction or knowledge-input obsolescence (δ) works in the opposite direction. The calibrated value of suppliers' surplus share ($\kappa/2$) is lower in more recent years, boosting growth just as Corollary 2 implies. Also according to Corollary 2, the increase in attraction cost convexity (χ_β) leads to a drop in average efficiency of input in use $\overline{\boldsymbol{\epsilon}_u}$, and growth slows.

7 Concluding Remarks

This paper is the first attempt to provide empirical evidences and a theory of the evolution dynamics of knowledge networks and their consequences. We draw new insights and observe new externalities with the dynamics *in* and *of* the network. In particular, the extended model highlights a *knowledge path externality* motivated by the spillover effect in the data after formation of pathways.

We emphasize that our model serves only as a stylized and tractable benchmark.

Therefore, we conclude this paper by pointing out a few directions of extensions for future research, especially enrichments for more quantitatively oriented studies. One is to consider richer network dynamics in this framework. For example, the search for new input sources may be path-dependent and thus directed by existing network connections, such that it is easier to search in nearby industries than to reach into more distant ones. Moreover, other extensive-margin dynamics can be introduced, such as the emergence and disappearance of industries, or firm births and deaths. Another natural extension is to allow for aggregate non-stationarity, which is useful to address issues associated with network resilience, the speed of recovery after negative shocks, and any structural change in innovation. Serious policy counterfactuals also call for appropriate modifications of the model.

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.

Acemoglu, Daron and Alireza Tahbaz-Salehi (2023): "The Macroeconomics of Supply Chain Disruptions," *NBER Working Paper*, 27565.

Aekka, Anuraag and Gaurav Khanna (2024): "Endogenous Production Networks and Firm Dynamics," Working paper.

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

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.

Arcidiacono, Peter, Patrick Bayer, Jason R. Blevins, and Paul B. Ellickson (2016): "Estimation of Dynamic Discrete Choice Models in Continuous

Time with an Application to Retail Competition," *The Review of Economic Studies*, 83 (3), 889–931.

Arkolakis, Costas, Federico Huneeus, and Yuhei Miyauchi (2023): *Spatial Production Networks*.

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.

Baqaee, David, Ariel Burstein, Cédric Duprez, and Emmanuel Farhi (2024): "Supplier Churn and Growth: A Micro-to-Macro Analysis," Working Paper 31231, National Bureau of Economic Research.

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

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.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101 (3), 471–475.

Demir, Banu, Ana Cecília Fieler, Daniel Yi Xu, and Kelly Kaili Yang (2024): "O-Ring Production Networks," *Journal of Political Economy*, 132 (1), 200–247.

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.

Eaton, Jonathan, Samuel S Kortum, and Francis Kramarz (2022): "Firm-to-firm trade: Imports, exports, and the labor market," Tech. rep., National Bureau of Economic Research.

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.

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

——— (2005): "Market value and patent citations," *RAND Journal of economics*, 16–38.

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.

Huneeus, Federico (2018): "Production network dynamics and the propagation of shocks," *Graduate thesis, Princeton University, Princeton, NJ*, 52.

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

Jordà, Òscar (2005): "Estimation and inference of impulse responses by local projections," *American economic review*, 95 (1), 161–182.

Jovanovic, Boyan and Peter L Rousseau (2005): "General purpose technologies," in *Handbook of economic growth*, Elsevier, vol. 1, 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.

Kopytov, Alexandr, Kristoffer Nimark, Bineet Mishra, and Mathieu Taschereau-Dumouchel (2023): "Endogenous Production Networks under Supply Chain Uncertainty," .

Lenoir, Clémence, Julien Martin, and Isabelle Mejean (2023): "Search frictions in international goods markets," *Journal of the European Economic Association*, 21 (1), 326–366.

Li, Fei, Charles Murry, Can Tian, and Yiyi Zhou (2024): "Inventory Management in Decentralized Markets," *International Economic Review*, 65 (1), 431–470.

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

Liu, Ernest and Song Ma (2024): "Innovation Networks and R&D Allocation," Working paper.

Miyauchi, Yuhei (2023): "Matching and Agglomeration: Theory and Evidence from Japanese Firm-to-Firm Trade," Working paper.

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

Panigrahi, Piyush (2021): "Endogenous spatial production networks: Quantitative implications for trade and productivity," .

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

Appendix

A Omitted Details and Proofs: Benchmark Model

A.1 Firm state and value in the benchmark model

Consider a firm of type (τ, ϵ) with relative productivity z at t. The profit flow is πz, where z declines due to aggregate growth at the endogenous $\boldsymbol{g}$. The upstream set is $U = \{\epsilon_u\}_{u \in U}$ and the downstream set is $D = \{(\tau_d, \mu_d)\}_{d \in D}$ with some notational abuse. Each link gets exogenously destroyed at δ. Each firm takes as given a cross-sectional joint distribution of firm states and decisions to infer $\boldsymbol{g}$ and the expected gain from searching or attracting. The stationary firm value function V obeys the following Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned}
\rho V(\tau, \epsilon, z, U, D) = \max_{\widehat{\mu}, \widehat{\lambda}, \widehat{\beta}} \Bigg\{ & \pi z - \boldsymbol{g} z v_z + \widehat{\mu}(1-b) S(\tau, \epsilon^U) - w\tau\epsilon^U c_\mu(\widehat{\mu}) + b \sum\nolimits_{d \in D} \mu_d S(\tau_d, \epsilon) \\
& + \widehat{\lambda}\, \mathbb{E}^{\widetilde{\beta}, \widetilde{\epsilon}} \left[\frac{\widetilde{\beta}}{\boldsymbol{\beta}} [V(\tau, \epsilon, z, U \cup \{\widetilde{\epsilon}\}, D) - V(\tau, \epsilon, z, U, D)] \right] - w c_\lambda(\widehat{\lambda}) \\
& + \frac{\widehat{\beta}}{\boldsymbol{\beta}}\, \mathbb{E}^{\widetilde{\lambda}, \widetilde{\mu}, \widetilde{\tau}} \left[\widetilde{\lambda} [V(\tau, \epsilon, z, U, D \cup \{(\widetilde{\tau}, \widetilde{\mu})\}) - V(\tau, \epsilon, z, U, D)] \right] - w c_\beta(\widehat{\beta}) \\
& + \delta \sum\nolimits_{u \in U} [V(\tau, \epsilon, z, U \setminus \{\epsilon_u\}, D) - V(\tau, \epsilon, z, U, D)] \\
& + \delta \sum\nolimits_{d \in D} [V(\tau, \epsilon, z, U, D \setminus \{(\tau_d, \mu_d)\}) - V(\tau, \epsilon, z, U, D)] \Bigg\}. \quad \text{(A.1)}
\end{aligned}$$

We solve for firm value at a stationary equilibrium by guess and verification. We guess a separable firm value such that $V(\tau, \epsilon, z, U, D) = z v_z + v^{up}(\tau, U) + v^{down}(\epsilon, D)$ consisting of value from the product market $z v_z$ and respective values associated with suppliers and users. We also guess that other firms' decisions are summarized by a common innovation rate $\boldsymbol{\mu}$, a search policy $\boldsymbol{\lambda} : \Omega \to \mathbb{R}_+$, and a attraction policy $\boldsymbol{\beta} : \Omega \to \mathbb{R}_+$. It is then sufficient to know the cross-sectional state distribution $\boldsymbol{M}$ to obtain $\mathbb{E}^{\widetilde{\beta}, \widetilde{\epsilon}}[\cdot]$ and $\mathbb{E}^{\widetilde{\lambda}, \widetilde{\mu}, \widetilde{\tau}}[\cdot]$. Then the HJB in Eq. (A.1) can be written in parts as Eqs. (2), (8) and (9) in the main text.

A.2 HJB equation in a generic setting

Generically, a firm's individual state shall include the information of *all walks passing through the firm*, that is, all firms that are directly and indirectly linked to it in any direction and at any arbitrary distance, such as each supplier's suppliers and each user's other suppliers and so on.

Let $\omega = (\tau, \epsilon, z, U, D)$ represent a firm 's individual state and Ω the set of all possible firm states. Firm value function is denoted as $V : \Omega \to \mathbb{R}_+$. The stationary cross-sectional measure over Ω is $\boldsymbol{M}$ such that $\boldsymbol{M}(\Omega) = 1$. In the benchmark model, the upstream set U and the downstream set D eventually reduce to type-weighted out- and indegrees, ϵ^U and τ^D. Generically, however, U and D should contain information of the individual states of all existing suppliers and users, including their links to other firms. In this case, $\omega = (\tau, \epsilon, z, U, D)$ may be recursively defined such that $U = \{\omega_u = (\tau_u, \epsilon_u, z_u, U_u, D_u)\}_{u \in U}$ and $D = \{\omega_d = (\tau_d, \epsilon_d, z_d, U_d, D_d)\}_{d \in D}$ at any time, where the indirected connected U_u, D_u, U_d, D_d have similar structures. Consequently, individual firm value may respond to changes in these higher-order states. In this case, we can heuristically write a generic HJB equation for firm value. A firm takes as given $\boldsymbol{M}$ and other firms' policies $\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta} : \Omega \to \mathbb{R}_+$ and infers aggregate growth $\boldsymbol{g}$. For brevity, the full list of firm state may be written as $\omega = (x, \omega_{-x})$ to highlight a particular element. Denote a generic innovation outcome as $z'(\omega)$ when a firm of state ω successfully innovates. In the benchmark model, $z'(\omega) = z + \tau\epsilon^U$. The endogenous innovation surplus becomes $S(\omega) = V(z'(\omega), \omega_{-z}) - V(\omega)$. The firm takes as given a surplus allocation rule (e.g., Shapley values) $S^U(\omega, \omega_u)$ that specifies the rent extracted from the innovating firm of state ω by an existing supplier $u \in U$ with state ω_u upon a successful innovation. In the benchmark model, the pairwise rent depends only on the corresponding types of the user-supplier pair involved, $S^U(\omega, \omega_u) = S^U(\tau, \epsilon_u) = bS(\tau, \epsilon_u) = b\tau\epsilon_u v_z$. Generically, the pairwise rent can be contingent on other suppliers' states. The HJB has a heuristic form as follows:

$$
\begin{aligned}
\rho V(\omega) =& \max_{\widehat{\mu}, \widehat{\lambda}, \widehat{\beta}} \Bigg\{ \pi z - \boldsymbol{g} z \frac{\partial V(\omega)}{\partial z} \\
&+ \widehat{\mu} \left(V(z'(\omega), \omega_{-z}) - V(\omega) - \sum\nolimits_{u \in U} S^U(\omega, \omega_u) \right) - w c_\mu(\widehat{\mu}, \omega) + \sum\nolimits_{d \in D} \boldsymbol{\mu}(\omega_d) S^U(\omega_d, \omega) \\
&+ \widehat{\lambda} \int \frac{\boldsymbol{\beta}(\widetilde{\omega})}{\int \boldsymbol{\beta}(\omega') d\boldsymbol{M}} [V(U \cup \{\widetilde{\omega}\}, \omega_{-U}) - V(\omega)] d\boldsymbol{M} - w c_\lambda(\widehat{\lambda})
\end{aligned}
$$

$$+\frac{\widehat{\beta}}{\int \boldsymbol{\beta}(\omega')d\boldsymbol{M}}\int \boldsymbol{\lambda}(\widetilde{\omega})[V(D\cup\{\widetilde{\omega}\},\omega_{-D})-V(\omega)]d\boldsymbol{M}-wc_{\beta}(\widehat{\beta})$$
$$+\delta\sum\nolimits_{u\in U}[V(U\setminus\{\omega_u\},\omega_{-U})-V(\omega)]+\delta\sum\nolimits_{d\in D}[V(D\setminus\{\omega_d\},\omega_{-D})-V(\omega)]$$
$$+\mathscr{D}V(\omega)\Big\}, \tag{A.2}$$

where the last term $\mathscr{D}V(\omega)$ captures the impact of higher-order state changes on firm value, accounting for all other firms that are k-step neighbors, $k=1,2,\ldots$, so it has indefinitely many components. The setup of the benchmark model ensures that $\mathscr{D}V(\omega)=0$.

A.2.1 What if the innovation outcome depends on individual z's?

The crucial assumption for $\mathscr{D}V(\omega)$ to have a finite "depth" is that the endogenous and network-dependent relative productivity z, a firm's own or its neighbors', is irrelevant for innovation or link formation incentives; otherwise, each firm needs to track the productivity evolution of all the directly and indirectly linked firms.

For example, suppose a firm's productivity jump upon a successful innovation at t depends on the suppliers' relative productivities $\{z_u\}_{u\in U}$ at the moment. It immediately means that the firm must know how each z_u evolves over time, which, in turn, relies on the evolving productivity levels of suppliers further upstream $\{\{z_{uu}\}_{uu\in U_u}\}_{u\in U}$ and so on. These dynamics affect the firm's incentives to search for new suppliers. Meanwhile, if surplus allocation also relies on $\{z_u\}_{u\in U}$, then the expected rent flows extracted from downstream users D depend on their $\{\{z_{ud}\}_{ud\in U_d}\}_{d\in D}$ at the moment. It follows that productivity evolution of direct and further-upstream suppliers of every user d affect a firm's incentives to attract new users. Each firm's individual decision problem becomes extremely complex if at all possible to be well-defined.

A.2.2 What if the suppliers are non-separable in a firm's payoff?

There are numerous ways to introduce supplier non-separability. An immediate consequence is that a firm's indirectly linked neighbors are payoff relevant, especially the other suppliers of each existing user. Therefore, $\mathscr{D}V(\omega)\neq 0$ in the HJB equation Eq. (A.2) for firm value. Moreover, the first moments of the cross-sectional in- and out-degree distributions, $\boldsymbol{x}$ or the weighted $\boldsymbol{X}$, are no longer sufficient statistics to summarize said distributions for firm value. Fortunately, $\mathscr{D}V(\omega)$ can still have a finite

depth, and the model remains (computationally) tractable.

We consider a small and simple deviation from the benchmark model to show that the model may accommodate supplier interactions. For better illustration, we remove type heterogeneity such that firms have identical types $\tau = a > 0$, and $\epsilon = 1$. Therefore, firms differ in endogenous states only. Consider a firm with $\omega = (z, U, D)$, where U is the set of existing suppliers, and D is the set of existing users. The relative productivity jump conditional on a successful innovation is simply ax^U, where $x^U = |U| \in N$ is the unweighted supplier count. Assume that the innovation cost in research labor has the form $C_\mu(\widehat{\mu}, x^U)$, where $\widehat{\mu}$ is the chosen innovation rate. Suppose $C_\mu(0, x^U) = 0$, $\partial C_\mu(\mu, x^U)/\partial\mu > 0$, and $\partial^2 C_\mu(\mu, x^U)/\partial\mu^2 > 0$, $\forall x^U \in \mathbb{N}$. The rest of the setup remains unchanged. In particular, surplus allocation remains unchanged.[12]

Given aggregate growth $\boldsymbol{g}$, we still have $v_z = \pi/(\rho+\boldsymbol{g})$. A firm's optimal innovation rate is now a function of its supplier count: $\boldsymbol{\mu}(x^U) = \arg\max_{\widehat{\mu}} \left\{\widehat{\mu}(1-b)ax^U v_z - wC_\mu(\widehat{\mu}, x^U)\right\}$, where $(1-b)$ remains the innovator's share of realized innovation surplus $ax^U v_z$ where every supplier makes an equal contribution av_z. Observe that the x^U suppliers are non-separable in the expected net benefit of innovation unless $C_\mu(\widehat{\mu}, x^U) = x^U c_\mu(\widehat{\mu})$ as in the benchmark case. It follows that the firm anticipates to extract from its users a flow rent of $bav_z \sum_{d\in D} \boldsymbol{\mu}(x_d^U)$. Such a small deviation from the benchmark leads to a few complications. First, individual firm state must include each user's supplier count x_d^U, $\forall d \in D$. Second, relatedly, an existing user gaining or losing a supplier has direct impact on a firm's value. Third, the incentives to search for new suppliers depend on the number of existing suppliers x^U.

Let $\boldsymbol{F}^U$ denote the stationary cross-sectional distribution of supplier count (out-degree) $\widetilde{x^U}$ in equilibrium with pmf $\boldsymbol{f}^U$. It will directly enter each firm's value and decision. Furthermore, growth rate becomes $\boldsymbol{g} = a\sum_{x^U=0}^{\infty} \boldsymbol{\mu}(x^U)x^U \boldsymbol{f}^U(x^U)$ when well-defined, which also requires the distribution $\boldsymbol{F}^U$. As will be shown, the equilibrium search policy is a function of supplier count, $x^U \mapsto \boldsymbol{\lambda}(x^U)$. When $\boldsymbol{\lambda}(\cdot)$ permits a

[12] We introduce supplier non-separability in the cost function for expositional simplicity. Alternatively, one may modify the production function and let the productivity jump be $a(x^U)$ for some function $a(\cdot)$, which complicates surplus allocation. Nonetheless, it does not affect the fact that each existing user's other suppliers are payoff relevant to any firm.

well-defined $\boldsymbol{F}^U$, the pmf is unique and satisfies

$$\begin{aligned}\boldsymbol{f}^U(x) =& \frac{\boldsymbol{\lambda}(x-1)}{\delta x}\boldsymbol{f}^U(x-1) = \frac{\boldsymbol{\lambda}(x-1)\times\ldots\times\boldsymbol{\lambda}(0)}{\delta^x x!}\boldsymbol{f}^U(0), \quad x\geq 1,\\ \boldsymbol{f}^U(0) =& \left(1+\sum\nolimits_{x=1}^{\infty}\frac{\boldsymbol{\lambda}(x-1)\times\ldots\times\boldsymbol{\lambda}(0)}{\delta^x x!}\right)^{-1}.\end{aligned}$$

Firms take this distribution as given when making decisions.

We proceed to show the system of HJB equations for firm value, whose form is conjectured to be $V(\tau,\epsilon,z,U,D) = zv_z + v^{up}(x^U) + v^{down}(D)$, where $D = \{x_d^U\}_{d\in D}$. Firms take as given some aggregate growth rate $\boldsymbol{g}$, the cross-sectional outdegree distribution $\boldsymbol{F}^U$, and others' policies $x^U \mapsto \boldsymbol{\mu}(x^U)$, $x^U \mapsto \boldsymbol{\lambda}(x^U)$, $D \mapsto \boldsymbol{\beta}(D)$. Then v_z follows. The HJB equation for $v^{up}(x^U)$ evaluated in equilibrium at $x^U \geq 1$ satisifes

$$\begin{aligned}\rho v^{up}(x^U) =& \boldsymbol{\mu}(x^U)(1-b)ax^U v_z - wC_\mu(\boldsymbol{\mu}(x^U),x^U)\\ &+\max\nolimits_{\widehat{\lambda}}\{\widehat{\lambda}[v^{up}(x^U+1)-v^{up}(x^U)] - wc_\lambda(\widehat{\lambda}) + \delta x^U[v^{up}(x^U-1)-v^{up}(x^U)]\},\end{aligned}$$

and at $x^U = 0$, $\rho v^{up}(0) = \max_{\widehat{\lambda}}\{\widehat{\lambda}[v^{up}(1)-v^{up}(0)] - wc_\lambda(\widehat{\lambda})\}$. Generically, the system cannot be analytically solved, but numerical computation is straightforward given v_z. So far, firm problem does not differ much from the benchmark.

The HJB equation for $v^{down}(D)$ can be written as

$$\begin{aligned}\rho v^{down}(D) =& \max\nolimits_{\widehat{\beta}}\Big\{\sum\nolimits_{d\in D}\boldsymbol{\mu}(x_d^U)b\tau_d v_z\\ &+\frac{\widehat{\beta}}{\boldsymbol{\beta}}\sum\nolimits_{\widetilde{x^U}=0}^{\infty}\boldsymbol{\lambda}(\widetilde{x^U})[v^{down}(D\cup\{\widetilde{x^U}\})-v^{down}(D)]\boldsymbol{f}^U(\widetilde{x^U}) - wc_\beta(\widehat{\beta})\\ &+\delta\sum\nolimits_{d\in D}[v^{down}(D\setminus\{x_d^U\})-v^{down}(D)]\\ &+\sum\nolimits_{d\in D}\boldsymbol{\lambda}(x_d^U)[v^{down}(\{x_d^U+1,D_{-d}\})-v^{down}(D)]\\ &+\delta\sum\nolimits_{d\in D}(x_d^U-1)[v^{down}(\{x_d^U-1,D_{-d}\})-v^{down}(D)]\Big\},\end{aligned}$$

where the last two lines describe the impact of each existing user gaining a new supplier or losing an existing one that is not the focal firm. These events are payoff relevant because of $\boldsymbol{\mu}(x_d^U)$ as the expected rent extracted from each user depends on the user's other suppliers. We further guess and verify that $v^{down}(D) = \sum_{d\in D}v_D(x_d^U) + v_{\Delta D}$, where $v_D(x_d^U)$ is the value of the individual link to an existing user d when d has x_d^U suppliers, and $v_{\Delta D}$ is the option value of attracting a new user.

Note that $x_d^U \geq 1$ always holds. We rewrite the HJB as

$$\rho v_D(x_d^U) = \boldsymbol{\mu}(x_d^U) b \tau_d v_z - \delta v_D(x_d^U) + \delta(x_d^U - 1)[v_D(x_d^U - 1) - v_D(x_d^U)] + \boldsymbol{\lambda}(x_d^U)[v_D(x_d^U + 1) - v_D(x_d^U)]$$

$$\rho v_{\Delta D} = \max_{\widehat{\beta}} \left\{ \frac{\widehat{\beta}}{\overline{\boldsymbol{\beta}}} \sum_{\widetilde{x^U}=0}^{\infty} \boldsymbol{\lambda}(\widetilde{x^U}) v_D(\widetilde{x^U}) \boldsymbol{f}^U(\widetilde{x^U}), -w c_\beta(\widehat{\beta}) \right\}.$$

Numerically solving for a stationary equilibrium in this case is relatively simple.

A.3 Proofs

Proof of Proposition 1. We begin with the unweighted average degree $\boldsymbol{x}$. Note that it is always true in directed networks that $\boldsymbol{x} = \overline{x^U} = \overline{x^D}$, where $x^U := |U|$ is a firm's supplier count and $x^D := |D|$ its user count at a given time. Stationarity requires balanced creation and destruction of links. In the aggregate, the flow of new links is deterministic: $\int \boldsymbol{\lambda}(\tau) dF_\tau(\tau) = \overline{\boldsymbol{\lambda}}$, and the flow of destroyed links is also deterministic at $\delta \boldsymbol{x}$. Therefore, $0 = \dot{\boldsymbol{x}} = \overline{\boldsymbol{\lambda}} - \delta \boldsymbol{x}$ implies $\boldsymbol{x} = \frac{\overline{\lambda}}{\delta}$.

We turn to the match-quality-weighted average degree $\boldsymbol{X}$. Similar to the case of unweighted degrees, aggregation of weighted links across *nodes* in either direction is equivalent to aggregation across *links* directly. That is, $\boldsymbol{X} = \int \tau \epsilon^U d\boldsymbol{M} = \int \tau^D \epsilon d\boldsymbol{M}$. From t to $t + dt$, a mass $\overline{\boldsymbol{\lambda}} dt$ of new links is created, such that a randomly drawn one is from a firm of output type no greater than τ to a firm of input type no greater than ϵ' with probability $H(\tau \to \epsilon') = \frac{1}{\overline{\boldsymbol{\lambda}} dt} \int_{\tau_0}^{\tau} \int_{\epsilon_0}^{\epsilon'} \boldsymbol{\lambda}(\widetilde{\tau}) \frac{\boldsymbol{\beta}(\widetilde{\epsilon}) dF_\tau dF_\epsilon}{\overline{\boldsymbol{\beta}}} dt = \frac{\int_{\tau_0}^{\tau} \boldsymbol{\lambda}(\widetilde{\tau}) dF_\tau}{\overline{\boldsymbol{\lambda}}} \times \frac{\int_{\epsilon_0}^{\epsilon'} \boldsymbol{\beta}(\widetilde{\epsilon}) dF_\epsilon}{\overline{\boldsymbol{\beta}}}$. Stationarity requires the probability distribution of new links' match-quality and that of all links' to be the same. It follows that $\boldsymbol{X} = \boldsymbol{x} \int \tau \epsilon' dH(\tau \to \epsilon') = \overline{\boldsymbol{\tau}_d} \boldsymbol{x} \overline{\boldsymbol{\epsilon}_u}$.

We then show $\overline{\boldsymbol{\tau}_d} > \overline{\tau}$ and $\overline{\boldsymbol{\epsilon}_u} > \overline{\epsilon}$. As will be shown in the proof of Proposition 2, equilibrium $\boldsymbol{\lambda}(\tau)$ and $\boldsymbol{\beta}(\epsilon)$ strictly increase. Without loss, assume that F_τ and F_ϵ have well-defined pdfs f_τ and f_ϵ. Observe that $f_{\tau_d}(\tau) := \frac{\boldsymbol{\lambda}(\tau)}{\overline{\boldsymbol{\lambda}}} f_\tau(\tau)$ and $f_{\epsilon_u}(\epsilon) := \frac{\boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}} f_\epsilon(\epsilon)$ are also well-defined pdfs. Observe also that $\frac{f_{\tau_d}(\tau)}{f_\tau(\tau)}$ and $\frac{f_{\epsilon_u}(\epsilon)}{f_\epsilon(\epsilon)}$ strictly increase by construction. This is the (strict) monotone likelihood ratio property (MLRP), which implies (strict) first-order stochastic dominance (FOSD), so $\overline{\boldsymbol{\tau}_d} > \overline{\tau}$ and $\overline{\boldsymbol{\epsilon}_u} > \overline{\epsilon}$.

Lastly, we show the expression for growth in Eq. (13). As will be shown in the proof of Proposition 2, equilibrium $\boldsymbol{\mu}$ is common and constant. By definition, growth $\boldsymbol{g}$ satisfies $\boldsymbol{g} = \frac{\dot{\boldsymbol{Z}}(t)}{\boldsymbol{Z}(t)} = \frac{1}{\boldsymbol{Z}(t)} \int \boldsymbol{\mu} \tau \epsilon^U \boldsymbol{Z}(t) d\boldsymbol{M} = \boldsymbol{\mu} \boldsymbol{X}$, and Eq. (13) follows. □

Proof of Proposition 2. We solve for the value function and find the associated

policies in equilibrium by guess and verification.

First, we take as given any $\boldsymbol{g} \geq 0$ and obtain $v_z = v_z(\boldsymbol{g}) > 0$ immediately. The surplus allocation rule follows. Then, regardless of firm state, the optimal innovation rate is given as $\mu^* = \mu^*(\boldsymbol{g}) = (c'_\mu)^{-1}\left((1-b)\frac{v_z}{w}\right) = \arg\max_{\widehat{\mu}}\{\widehat{\mu}(1-b)v_z - wc_\mu(\widehat{\mu})\}$. In equilibrium, $\boldsymbol{\mu} = \mu^*$, so it is without loss that firms take as given a common $\boldsymbol{\mu}$.

Second, we take as given a $\boldsymbol{\mu}$, a (non-degenerate) stationary cross-sectional distribution $\boldsymbol{M}$ and any policy functions $\boldsymbol{\lambda}, \boldsymbol{\beta} : \Omega \to \mathbb{R}_+$. It follows that $\overline{\boldsymbol{\beta}} = \int \boldsymbol{\beta}(\widetilde{\omega})d\boldsymbol{M}$ is a constant. We evaluate Eq. (8) at $\mu^* = \boldsymbol{\mu}$ and guess $v^{up}(\tau, \epsilon^U) = \tau\epsilon^U v_U + v_{\Delta U}(\tau)$ and $v^{down}(\epsilon, \tau^D) = \tau^D \epsilon v_D + v_{\Delta D}(\epsilon)$. Eqs. (8) and (9) become

$$\rho\tau\epsilon^U v_U = \tau\epsilon^U\left(\boldsymbol{\mu}(1-b)v_z - wc_\mu(\boldsymbol{\mu})\right) - \delta\tau\epsilon^U v_U, \quad \rho\tau^D\epsilon v_D = \boldsymbol{\mu}b\tau^D\epsilon v_z - \delta\tau^D\epsilon v_D,$$

$$\rho v_{\Delta U}(\tau) = \max_{\widehat{\lambda}}\left\{\widehat{\lambda}\int\frac{\boldsymbol{\beta}(\widetilde{\omega})}{\overline{\boldsymbol{\beta}}}\tau\widetilde{\epsilon}v_U d\boldsymbol{M} - wc_\lambda(\widehat{\lambda})\right\}, \quad \rho v_{\Delta D}(\epsilon) = \max_{\widehat{\beta}}\left\{\frac{\widehat{\beta}}{\overline{\boldsymbol{\beta}}}\int\boldsymbol{\lambda}(\widetilde{\omega})\widetilde{\tau}\epsilon v_D d\boldsymbol{M} - wc_\beta(\widehat{\beta})\right\}.$$

The system implies the expressions of v_U and v_D as endogenous coefficients. Moreover, it implies that the "best response" functions to the given $\boldsymbol{\lambda}(\cdot)$ and $\boldsymbol{\beta}(\cdot)$ are simply $\lambda^*(\tau) = \lambda^*(\tau; \boldsymbol{g}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{M}) = (c'_\lambda)^{-1}\left(\tau\frac{\int\widetilde{\epsilon}\boldsymbol{\beta}(\widetilde{\omega})d\boldsymbol{M}}{\overline{\boldsymbol{\beta}}}\frac{v_U}{w}\right)$, and $\beta^*(\epsilon) = \beta^*(\tau; \boldsymbol{g}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{M}) = (c'_\beta)^{-1}\left(\frac{\epsilon}{\overline{\boldsymbol{\beta}}}\frac{v_D}{w}\int\widetilde{\tau}\boldsymbol{\lambda}(\widetilde{\omega})d\boldsymbol{M}\right)$, which depend on firm state via respective types only. Therefore, the equilibrium policies $\boldsymbol{\lambda}(\cdot)$ and $\boldsymbol{\beta}(\cdot)$ must be functions of respective types only. The given statistics are sufficient information of the distribution $\boldsymbol{M}$. Expressions of $v_{\Delta U}(\tau)$ and $v_{\Delta D}(\epsilon)$ are obtained at $\lambda^*(\tau) = \boldsymbol{\lambda}(\tau)$ and $\beta^*(\epsilon) = \boldsymbol{\beta}(\epsilon)$. The convexity of $v_{\Delta U}(\tau)$ and $v_{\Delta D}(\epsilon)$ follows by applying the Envelop Theorem.

Lastly, an equilibrium requires $\boldsymbol{g}$ to be consistent with the distribution and policies, which is Eq. (13). □

Proof of Proposition 3. We begin by showing the uniqueness of a stationary equilibrium that is fully described by $\boldsymbol{g}$, $(\boldsymbol{\mu}, \boldsymbol{\lambda}(\cdot), \boldsymbol{\beta}(\cdot))$, and (v_z, v_U, v_D) that satisfy the conditions listed in Proposition 2. Observe that at $\boldsymbol{g} = g \geq 0$, (v_z, v_U, v_D) and $\boldsymbol{\mu}$ are uniquely pinned down, and they strictly decrease in g. Let $\boldsymbol{\mu}$ be a function of g denoted as $\boldsymbol{\mu}(g)$. We proceed to show that a given g also uniquely pins down $\boldsymbol{\lambda}(\cdot)$ and $\boldsymbol{\beta}(\cdot)$. Using the assumed functional forms, we have $\boldsymbol{\lambda}(\tau) \propto \left(\tau\overline{\boldsymbol{\epsilon}_u}\frac{v_U}{w}\right)^{1/\chi_\lambda} \propto \tau^{1/\chi_\lambda}$, and $\boldsymbol{\beta}(\epsilon) \propto \left(\frac{\epsilon}{\overline{\boldsymbol{\beta}}}\overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}\frac{v_D}{w}\right)^{1/\chi_\beta} \propto \epsilon^{1/\chi_\beta}$. It follows that $\frac{\boldsymbol{\lambda}(\tau)}{\overline{\boldsymbol{\lambda}}} = \frac{\tau^{1/\chi_\lambda}}{\overline{\tau^{1/\chi_\lambda}}}$, and $\frac{\boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}} = \frac{\epsilon^{1/\chi_\beta}}{\overline{\epsilon^{1/\chi_\beta}}}$. The expressions of $\overline{\boldsymbol{\tau}_d}$ and $\overline{\boldsymbol{\epsilon}_u}$ are immediate and unique, and they are independent of other endogenous variables. It follows that $\boldsymbol{\lambda}(\tau)$ is point-wise pinned down by v_U, which is

in turn pinned down by g. Furthermore, it is easy to see that $\boldsymbol{\lambda}(\tau)$ strictly decreases in g at any τ. Then, $\overline{\boldsymbol{\lambda}}$ follows as a strictly decreasing function of g denoted as $\overline{\boldsymbol{\lambda}}(g)$, which yields a unique function $\boldsymbol{\beta}(\epsilon)$. Lastly, it suffices to show that a unique $\boldsymbol{g}$ satisfies the equilibrium conditions. Define a mapping $\mathcal{T}$ such that $\mathcal{T}(g) = \boldsymbol{\mu}(g)\overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}(g)\overline{\boldsymbol{\epsilon}_u}/\delta$. Therefore, $\mathcal{T}(g)$ is continuous and strictly decreases in g and has a unique fixed point $\boldsymbol{g} = \mathcal{T}(\boldsymbol{g}) \in (0, \mathcal{T}(0))$, implying uniqueness of a stationary equilibrium.

We proceed to show the properties. First, we show that $\overline{\boldsymbol{\tau}_d}$ and $\overline{\boldsymbol{\epsilon}_u}$ decrease in χ_λ and χ_β, respectively. Without loss, assume well-defined pdfs f_τ and f_ϵ. Any given $\chi_\lambda > 0$ and $\chi_\beta > 0$ define two other pdfs that result in $\overline{\boldsymbol{\tau}_d}$ and $\overline{\boldsymbol{\epsilon}_u}$: $\frac{\boldsymbol{\lambda}(\tau)}{\overline{\boldsymbol{\lambda}}} f_\tau(\tau) = \frac{\tau^{1/\chi_\lambda}}{\overline{\tau^{1/\chi_\lambda}}} f_\tau(\tau)$, and $\frac{\boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}} f_\epsilon(\epsilon) = \frac{\epsilon^{1/\chi_\beta}}{\overline{\epsilon^{1/\chi_\beta}}} f_\epsilon(\epsilon)$. Suppose $\chi_{\lambda,1} < \chi_{\lambda,2}$, then the density ratio $\frac{\tau^{1/\chi_{\lambda,1}}}{\overline{\tau^{1/\chi_{\lambda,1}}}} f_\tau(\tau) / \frac{\tau^{1/\chi_{\lambda,2}}}{\overline{\tau^{1/\chi_{\lambda,2}}}} f_\tau(\tau) = \tau^{1/\chi_{\lambda,1} - 1/\chi_{\lambda,2}} \overline{\tau^{1/\chi_{\lambda,2}}} / \overline{\tau^{1/\chi_{\lambda,1}}}$ strictly increases in τ, which is MLRP and implies FOSD. It follows that $\overline{\boldsymbol{\tau}_d}$ decreases in χ_λ. It is similar to show that $\overline{\boldsymbol{\epsilon}_u}$ decreases in χ_β.

Second, we show Eq. (18). The assumed functional form implies $\boldsymbol{\mu} \propto \left((1-b)\frac{v_z}{w}\right)^{1/\chi_\mu}$, so $v_U \propto \frac{1}{\rho+\delta} \frac{((1-b)v_z)^{1+1/\chi_\mu}}{w^{1/\chi_\mu}} = \frac{1}{\rho+\delta} \frac{((1-b)\pi)^{1+1/\chi_\mu}}{(\rho+\boldsymbol{g})^{1+1/\chi_\mu} w^{1/\chi_\mu}}$. In equilibrium, we have

$$\begin{aligned}
\frac{d\ln \boldsymbol{g}}{d\ln \pi} =& \frac{d\ln \boldsymbol{\mu}}{d\ln \pi} + \frac{d\ln \overline{\boldsymbol{\tau}_d}}{d\ln \pi} + \frac{d\ln \boldsymbol{x}}{d\ln \pi} + \frac{d\ln \overline{\boldsymbol{\epsilon}_u}}{d\ln \pi} = \frac{1}{\chi_\mu}\frac{d\ln v_z}{d\ln \pi} + 0 + \frac{d\ln \overline{\boldsymbol{\lambda}}}{d\ln \pi} + 0 \\
=& \frac{1}{\chi_\mu}\left(1 - \frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}\frac{d\ln \boldsymbol{g}}{d\ln \pi}\right) + \frac{1}{\chi_\lambda}\frac{d\ln v_U}{d\ln \pi} \\
=& \frac{1}{\chi_\mu}\left(1 - \frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}\frac{d\ln \boldsymbol{g}}{d\ln \pi}\right) + \frac{1}{\chi_\lambda}\frac{\chi_\mu+1}{\chi_\mu}\left(1 - \frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}\frac{d\ln \boldsymbol{g}}{d\ln \pi}\right).
\end{aligned}$$

Rearranging yields the expression: $\frac{d\ln \boldsymbol{g}}{d\ln \pi} = \frac{\frac{1}{\chi_\mu} + \frac{1}{\chi_\lambda}\frac{\chi_\mu+1}{\chi_\mu}}{1+\left(\frac{1}{\chi_\mu} + \frac{1}{\chi_\lambda}\frac{\chi_\mu+1}{\chi_\mu}\right)\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}} > \left.\frac{d\ln \boldsymbol{g}}{d\ln \pi}\right|_{\text{exog. net.}} = \frac{\frac{1}{\chi_\mu}}{1+\frac{\chi_\mu+1}{\chi_\mu}\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}} > 0$, where $\left.\frac{d\ln \boldsymbol{g}}{d\ln \pi}\right|_{\text{exog. net.}}$ ignores the amplification effect of the endogenous network while accounting for the equilibrium effect via $\boldsymbol{g}$. The cases of w and $(1-b)$ are analogous and omitted.

Third, we show Eq. (19). As in the previous step, we have

$$\begin{aligned}
\frac{d\ln \boldsymbol{g}}{d\ln \delta} =& \frac{d\ln \boldsymbol{\mu}}{d\ln \delta} + 0 + \frac{d\ln \overline{\boldsymbol{\lambda}}}{d\ln \delta} - 1 + 0 \\
=& -\frac{1}{\chi_\mu}\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}\frac{d\ln \boldsymbol{g}}{d\ln \delta} + \frac{1}{\chi_\lambda}\left(-\frac{\delta}{\rho+\delta} - \frac{\chi_\mu+1}{\chi_\mu}\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}\frac{d\ln \boldsymbol{g}}{d\ln \delta}\right) - 1,
\end{aligned}$$

where the -1 on the right-hand side is the direct effect of δ on $\boldsymbol{x}$ and $\boldsymbol{g}$. Rearranging

yields $\frac{d\ln \boldsymbol{g}}{d\ln \delta} = -\frac{1+\frac{1}{\chi_\lambda}\frac{\delta}{\rho+\delta}}{1+\left(\frac{1}{\chi_\mu}+\frac{1}{\chi_\lambda}\frac{\chi_\mu+1}{\chi_\mu}\right)\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}} < 0$, which depends on the value of δ. In contrast, if the network is exogenous such that $\overline{\boldsymbol{\lambda}}$ is kept fixed, we have $\left.\frac{d\ln \boldsymbol{g}}{d\ln \delta}\right|_{\text{exog. net.}} = -\frac{1}{1+\frac{1}{\chi_\mu}\frac{\boldsymbol{g}}{\rho+\boldsymbol{g}}} < 0$. It means that, as $\delta \to 0$, and $\boldsymbol{g}$ gets large, $\left|\frac{d\ln \boldsymbol{g}}{d\ln \delta}\right| < \left|\frac{d\ln \boldsymbol{g}}{d\ln \delta}|_{\text{fixed net.}}\right| < 1$, the formation endogeneity mitigates the negative impact on growth caused by increased rate of knowledge obsolescence. However as $\delta \to \infty$, and $\boldsymbol{g}$ gets small, such endogeneity can worsen the negative impact, with $\left|\frac{d\ln \boldsymbol{g}}{d\ln \delta}\right| > 1 > \left|\frac{d\ln \boldsymbol{g}}{d\ln \delta}|_{\text{fixed net.}}\right|$. □

Proof of Proposition 4. It is obvious that a firm's supplier count x^U evolves as a birth-death process: x^U becomes x^U+1 at Poisson (birth) rate $\boldsymbol{\lambda}(\tau)$, $\forall x^U \in \mathbb{N}$, and x^U becomes x^U-1 at Poisson (death) rate δx^U if $x^U \geq 1$. The process for its user count x^D is similar, with birth rate $\overline{\boldsymbol{\lambda}}\boldsymbol{\beta}(\epsilon)/\overline{\boldsymbol{\beta}}$ and death rate δx^D.

We start from the outdegree or supplier count (number of pathways found) x^U, using the properties of the birth-death process. Its ergodic distribution has a probability mass function (pmf) that satisfies

$$\begin{aligned}\boldsymbol{f}^U(x \mid \tau) =& \frac{\boldsymbol{\lambda}(\tau)}{\delta x}\boldsymbol{f}^U(x-1 \mid \tau) = \left(\frac{\boldsymbol{\lambda}(\tau)}{\delta}\right)^x \frac{1}{x!}\boldsymbol{f}^U(0 \mid \tau),\\ \boldsymbol{f}^U(0 \mid \tau) =& \left(\sum\nolimits_{x=0}^{\infty}\left(\frac{\boldsymbol{\lambda}(\tau)}{\delta}\right)^x \frac{1}{x!}\right)^{-1} = e^{-\boldsymbol{\lambda}(\tau)/\delta},\end{aligned}$$

which is the pmf of a Poisson distribution with mean $\boldsymbol{\lambda}(\tau)/\delta = \boldsymbol{x}\boldsymbol{\lambda}(\tau)/\overline{\boldsymbol{\lambda}}$. The indegree distribution is analogous. □

Proof of Corollary 1. The expressions for $\mathbb{E}[\epsilon^U \mid \tau]$ and $\mathbb{E}[\tau^D \mid \epsilon]$ immediately follow Proposition 4. Given τ, ϵ^U, and z at t, the law of motion for z at the moment is $dz = -\boldsymbol{g}z dt + \tau\epsilon^U d\mathrm{P}(\boldsymbol{\mu}, t)$, where $\mathrm{P}(\boldsymbol{\mu}, t)$ represents the Poisson process for innovation at rate $\boldsymbol{\mu}$. Aggregating the law of motion within firms of type τ yields $\mathbb{E}[dz \mid \tau] = -\boldsymbol{g}\mathbb{E}[z \mid \tau]dt + \mathbb{E}[\tau\epsilon^U \mid \tau]\boldsymbol{\mu}dt$, so stationarity requires $0 = -\boldsymbol{g}\mathbb{E}[z \mid \tau] + \mathbb{E}[\tau\epsilon^U \mid \tau]\boldsymbol{\mu}$. Rearranging and plugging in $\boldsymbol{g} = \boldsymbol{\mu}\boldsymbol{X}$ yields the expression $\mathbb{E}[z \mid \tau] = \mathbb{E}[\tau\epsilon^U \mid \tau]\boldsymbol{\mu}/\boldsymbol{g} = \mathbb{E}[\tau\epsilon^U \mid \tau]/\boldsymbol{X} = \tau\boldsymbol{x}\boldsymbol{\lambda}(\tau)\overline{\boldsymbol{\epsilon}_u}/(\overline{\boldsymbol{\tau}_d}\boldsymbol{x}\overline{\boldsymbol{\epsilon}_u}\overline{\boldsymbol{\lambda}}) = \tau\boldsymbol{\lambda}(\tau)/(\overline{\boldsymbol{\tau}_d}\cdot\overline{\boldsymbol{\lambda}})$. □

B Omitted Details and Proofs: Extended Model

B.1 Firm state, firm value, and equilibrium definition

Firm states. Generically, each firm (j, f)'s individual state at t is summarized by $\omega_{jf}(t) = \omega = (\tau, \epsilon, z, J, U, D)$. It includes the firm's fixed type $\tau_{jf} = \tau$, $\epsilon_j = \epsilon$, and its relative productivity $z_{jf}(t) = z$ at t. J is the list of all firms' states in j, i.e., $J = (\omega_{jf'}(t))_{f' \in j}$. Sets $U_j(t) = U$ and $D_j(t) = D$ include the upstream and downstream industries, which can be recursively defined to include *all walks passing through the firm*, i.e., all of the firm's direct and indirect connections in any direction and of arbitrary lengths. The set of all possible firm states is Ω. Henceforth, we refer to a firm interchangeably by its index (j, f) or f and its state ω.

Thanks to the simplifying assumptions, firm f's state will eventually reduce to $\omega = (\tau, \tau_{-f}, \epsilon, z, \epsilon^U, \tau^D)$, where τ is the firm's own output type, τ_{-f} contains the output types of the other $n-1$ firms within the industry, $\epsilon^U = \sum_{u \in U} \epsilon_u$ is the combined supplier type for U, and $\tau^D = \frac{1}{n} \sum_{d \in D} \sum_{f' \in d} \tau_{df'}$ is the combined user type after adjusting for the probability $1/n$ of firm f being the matched supplier. Indeed, we can write ϵ^U in the analogous form such that $\epsilon^U = \sum_{u \in U} \epsilon_u = \sum_{u \in U} \frac{1}{n_u} \sum_{f' \in u} \epsilon_u$, where each of the n_u firms $f' \in u$ with the same ϵ_u becomes the matched supplier to f at the next successful innovation with probability $1/n_u$.

Aggregate state. Along the equilibrium path, each firm takes as given the time-invariant aggregate state $\boldsymbol{A} \equiv (\mathbf{w}, \boldsymbol{\pi}, \boldsymbol{\mu}, S^U, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{M})$. Here, $\mathbf{w} > 0$ is the constant researcher wage rate, $\boldsymbol{\pi}$ is the average profit, $\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta} : \Omega \to \mathbb{R}_+$ are other firms' innovation, search, and attraction policies, respectively. The supplier surplus $S^U : \Omega^2 \to \mathbb{R}_+$ characterizes surplus allocation as Shapley values. $\boldsymbol{M}$ is the stationary cross-sectional measure over firm states Ω with $\boldsymbol{M}(\Omega) = N$. The equilibrium law of motion of each firm's relative productivity z at t is $dz = -\boldsymbol{g} z dt + \tau(\epsilon^U + \eta\epsilon) d\mathrm{P}(\boldsymbol{\mu}, t) + \left(\int \frac{\boldsymbol{\beta}(\widetilde{\omega})}{\int \boldsymbol{\beta}(\omega') d\boldsymbol{M}} \frac{\gamma}{\widetilde{n}} \tau \widetilde{\epsilon} d\boldsymbol{M} \right) d\mathrm{P}(\boldsymbol{\lambda}, t)$, where $\boldsymbol{g}$ is the growth rate in average productivity, and $\mathrm{P}(\cdot, t)$ represents a Poisson jump process of a given rate at t. Firms infer from $\boldsymbol{A}$ the productivity growth rate $\boldsymbol{g} = \boldsymbol{g}(\boldsymbol{A})$ as

$$\boldsymbol{g} = \frac{\dot{\boldsymbol{Z}}(t)}{\boldsymbol{Z}(t)} = \frac{1}{N} \left(\boldsymbol{\mu} \int \tau(\epsilon^U + \eta\epsilon) d\boldsymbol{M} + \iint \boldsymbol{\lambda}(\omega) \frac{\boldsymbol{\beta}(\widetilde{\omega})}{\int \boldsymbol{\beta}(\omega') d\boldsymbol{M}} \frac{\gamma}{\widetilde{n}} \tau \widetilde{\epsilon} d\boldsymbol{M} d\boldsymbol{M} \right). \quad \text{(B.1)}$$

Firm value. Given $\boldsymbol{A}$, the stationary firm value is $V(\cdot; \boldsymbol{A}) : \Omega \to \mathbb{R}$ for every

firm. When not confusing, we drop $\boldsymbol{A}$ and write firm value at state ω as $V(\omega)$. It is useful to show formally the surplus allocation rule S^U. The proof is in Appendix B.

Lemma B.1. *Suppose firm value in equilibrium has the form $V(\omega) = zv_z + v(\omega_{-z})$, $v_z > 0$. Let $S^U(\tau, \epsilon') = \frac{\kappa}{2}\tau\epsilon' v_z$, $\forall \tau, \epsilon'$. Upon a successful innovation by a firm of output type τ using matched upstream suppliers of combined input type ϵ^U and a matched supplier in own industry of input type ϵ, the Shapley value of each matched supplier in $u \in U$ is $S^U(\tau, \epsilon_u) = \frac{\kappa}{2}\tau\epsilon_u v_z$, and that of the matched supplier in own industry is $S^U(\tau, \eta\epsilon) = \frac{\kappa}{2}\eta\tau\epsilon_u v_z$. The innovator gets the remaining surplus. At the formation of a new link from a user (pathfinder) of output type τ to a supplier (receiver) in an industry of input type ϵ' with n' firms, the supplier (receiver) gets $\boldsymbol{S}^U(\tau, \gamma\epsilon'/n') = \frac{\kappa}{2}\frac{\gamma}{n'}\tau\epsilon' v_z$, and the finder gets the remaining surplus.*

Consider a particular firm f in industry j. A more general HJB equation for firm value $V(\omega) = zv_z + v(\omega_{-z})$ can be heuristically written as follows.

$$
\begin{aligned}
\rho V(\omega) = \max_{\widehat{\mu},\widehat{\lambda},\widehat{\beta}} \Bigg\{ & \boldsymbol{\pi} z - \mathbf{w}\left(\tau(\epsilon^U + \eta\epsilon)c_\mu(\widehat{\mu}) + c_\lambda(\widehat{\lambda}) + \frac{1}{n}c_\beta(\widehat{\beta})\right) - \boldsymbol{g} z v_z \\
& + \underset{\text{innov.}}{\widehat{\mu}} \times \underbrace{\left(\tau(\epsilon^U + \eta\epsilon)v_z - \sum\nolimits_{u\in U} S^U(\tau,\epsilon') - S^U(\tau,\eta\epsilon)\right)}_{\text{sharing surplus with the randomly matched supplier in each industry}} \\
& + \underset{\text{search}}{\widehat{\lambda}} \times \underbrace{\int \frac{\boldsymbol{\beta}(\widetilde{\omega})}{\int \boldsymbol{\beta}(\omega')d\boldsymbol{M}} \left(\frac{\gamma}{\widetilde{n}}\tau\widetilde{\epsilon}v_z - S^U(\tau,\gamma\widetilde{\epsilon}/\widetilde{n}) + V(U\cup\{\widetilde{j}\},\omega_{-U}) - V(\omega)\right) d\boldsymbol{M}}_{\text{finding a new supplier or pathway}} \\
& + \underbrace{\sum\nolimits_{f'\in j\setminus\{f\}} \boldsymbol{\lambda}(\omega_{f'}) \int \frac{\boldsymbol{\beta}(\widetilde{\omega})}{\int \boldsymbol{\beta}(\omega')d\boldsymbol{M}} \left(V(U\cup\{\widetilde{j}\},\omega_{-U}) - V(\omega)\right) d\boldsymbol{M}}_{\text{another firm in } J\text{, i.e., peer, finding a new supplier pathway}} \qquad \text{(PS)} \\
& + \underbrace{\frac{1}{n-1}\sum\nolimits_{f'\in j\setminus\{f\}} \boldsymbol{\mu}(\omega_{f'}) S^U(\tau',\eta\epsilon)}_{\text{exp'd surplus received from others in } J} + \underbrace{\frac{1}{n}\sum\nolimits_{d\in D}\sum\nolimits_{f'\in d} \boldsymbol{\mu}(\omega_{f'}) S^U(\tau_k,\epsilon)}_{\text{exp'd surplus received from firms in industries in } D} \\
& + \underbrace{\frac{\widehat{\beta}}{\int \boldsymbol{\beta}(\omega')d\boldsymbol{M}}}_{\text{attract}} \underbrace{\int \boldsymbol{\lambda}(\widetilde{\omega}) \left(S^U(\widetilde{\tau},\gamma\epsilon/n) + V(D\cup\{\widetilde{j}\},\omega_{-D}) - V(\omega)\right) d\boldsymbol{M}}_{\text{getting a new user, i.e., receiving a new pathway}} \\
& + \underbrace{\sum\nolimits_{f'\in j\setminus\{f\}} \frac{\boldsymbol{\beta}(\omega_{f'})}{\int \boldsymbol{\beta}(\omega'')d\boldsymbol{M}} \int \boldsymbol{\lambda}(\widetilde{\omega}) \left(V(D\cup\{\widetilde{j}\},\omega_{-D}) - V(\omega)\right) d\boldsymbol{M}}_{\text{another firms in } J\text{, i.e., peer, attracting a new user or receiving a new pathway}} \qquad \text{(PA)} \\
& + \underbrace{\delta \times \left(\sum\nolimits_{u\in U}[V(U\setminus\{u\},\omega_{-U}) - V(\omega)] + \sum\nolimits_{d\in D}[V(D\setminus\{d\},\omega_{-D}) - V(\omega)]\right)}_{\text{exogenous destruction of existing connections}}
\end{aligned}
$$

$$+ \underbrace{\mathscr{D}V(\omega)}_{\text{higher-order state changes}} \Bigg\}.$$

Higher-order state changes do not affect $V(\cdot)$, so $\mathscr{D}V(\omega) = 0$.

Compared with the benchmark model, the critical changes to the firm value's HJB equation are the components (PS) and (PA) capturing how firms "free ride" peer firms' effort to search for suppliers and attract users. We solve for the firm value function and examine the equilibrium policies in the upcoming section.

B.2 Characterizing the equilibrium

Definition 2. A *stationary equilibrium* is a list $\boldsymbol{A} = (\mathbf{w}, \boldsymbol{\pi}, \boldsymbol{\mu}, S^U, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{M})$ together with an aggregate productivity growth rate $\boldsymbol{g}$ and firm value function $V(\cdot, \boldsymbol{A}) : \Omega \to \mathbb{R}_+$ such that $\boldsymbol{\pi} = \frac{1}{N}\frac{L}{\nu - 1}$ results from household optimization, firm's optimal hiring, production, and product-pricing decisions, and the cleared production labor market; given $\boldsymbol{A}$, each firm infers $\boldsymbol{g}$ according to Eq. (B.1) and optimizes to achieve value $V(\omega; \boldsymbol{A})$, with associated policies $\mu^*(\omega) = \boldsymbol{\mu}(\omega)$, $\lambda^*(\omega) = \boldsymbol{\lambda}(\omega)$, $\beta^*(\omega) = \boldsymbol{\beta}(\omega)$, $\forall \omega \in \Omega$, and an implied surplus allocation rule that coincides with S^U; $\boldsymbol{M}$ and policies are consistent; researcher wage $\mathbf{w}$ clears the market for research labor.

B.2.1 Firm value and policies in equilibrium

We begin the analysis by simplifying firm value. Without loss, we look for an equilibrium with a constant and common innovation rate $\boldsymbol{\mu} > 0$ and type-specific search and attraction policies $\tau \mapsto \boldsymbol{\lambda}(\tau)$ and $\epsilon \mapsto \boldsymbol{\beta}(\epsilon)$. We guess and verify that (i) firm state is fully summarized by $\omega = (\tau, \tau_{-f}, \epsilon, z, \epsilon^U, \tau^D)$, and (ii) firm value takes the form $V(\omega) = zv_z + v^{in}(\tau, \tau_{-f}, \epsilon) + v^{up}(\tau, \tau_{-f}, \epsilon^U) + v^{down}(\epsilon, n, \tau^D)$, where n is the number of firms in own industry, i.e., the vector length of (τ, τ_{-f}). Here, $v^{in}(\cdot)$ represents the value concerning within-industry surplus sharing; although $v^{up}(\cdot)$ and $v^{down}(\cdot)$ still capture firm value of network origins, they reflect the knowledge path externality: firms benefit from new links formed by peer firms.

As in the benchmark model, it is useful to introduce the following moments. The aggregate attraction effort is $\int \boldsymbol{\beta}(\epsilon) d\boldsymbol{M} = N \int \boldsymbol{\beta}(\epsilon) dF_\epsilon = N\overline{\boldsymbol{\beta}}$, where $\overline{\boldsymbol{\beta}}$ is the average attraction effort per firm. The aggregate search rate is $\int \boldsymbol{\lambda}(\tau) d\boldsymbol{M} = N \int \boldsymbol{\lambda}(\tau) dF_\tau = N\overline{\boldsymbol{\lambda}}$, where $\overline{\boldsymbol{\lambda}}$ is the average search rate per firm. Furthermore, the mean output type

of pathfinder firms is $\overline{\boldsymbol{\tau}_d} = \frac{1}{N}\int \frac{\tau \boldsymbol{\lambda}(\tau)}{\overline{\boldsymbol{\lambda}}} d\boldsymbol{M} = \int \frac{\tau \boldsymbol{\lambda}(\tau)}{\overline{\boldsymbol{\lambda}}} dF_\tau$, and the mean input type of path-receivers is $\overline{\boldsymbol{\epsilon}_u} = \frac{1}{N}\int \frac{\epsilon \boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}} d\boldsymbol{M} = \int \frac{\epsilon \boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}} dF_\epsilon$.

Given some $\boldsymbol{g} > 0$, $v_z = \frac{\boldsymbol{\pi}}{\rho + \boldsymbol{g}}$ follows immediately. Observe that, in equilibrium,

$$\mu^* = \boldsymbol{\mu} = \arg\max_{\widehat{\mu}} \left\{ \tau(\epsilon^U + \eta\epsilon) \left(\widehat{\mu}(1 - \frac{\kappa}{2}) v_z - \mathbf{w} c_\mu(\widehat{\mu}) \right) \right\}. \tag{B.2}$$

The rest of the system of HJB equations evaluated at $\mu^* = \boldsymbol{\mu}$ becomes

$$\begin{aligned}
&\rho v^{in}(\tau, \tau_{-f}, \epsilon) = \eta \left(\tau\epsilon \left(\boldsymbol{\mu}(1 - \frac{\kappa}{2}) v_z - \mathbf{w} c_\mu(\boldsymbol{\mu}) \right) + \boldsymbol{\mu} \frac{\kappa}{2} \frac{\sum_{\tau' \in \tau_{-f}} \tau'}{n-1} \epsilon v_z \right), \\
&\rho v^{up}(\tau, \tau_{-f}, \epsilon^U) \\
=& \max_{\widehat{\lambda}} \left\{ \tau\epsilon^U \left(\boldsymbol{\mu}(1 - \frac{\kappa}{2}) v_z - \mathbf{w} c_\mu(\boldsymbol{\mu}) \right) - \mathbf{w} c_\lambda(\widehat{\lambda}) + \delta \sum\nolimits_{u \in U} [v^{up}(\tau, \tau_{-f}, \epsilon^U - \epsilon_u) - v^{up}(\tau, \tau_{-f}, \epsilon^U)] \right. \\
&+ \widehat{\lambda} \int \frac{\boldsymbol{\beta}(\widetilde{\epsilon})}{N\overline{\boldsymbol{\beta}}} \left((1 - \frac{\kappa}{2}) \frac{\gamma}{\widetilde{n}} \tau \widetilde{\epsilon} v_z + v^{up}(\tau, \tau_{-f}, \epsilon^U + \widetilde{\epsilon}) - v^{up}(\tau, \tau_{-f}, \epsilon^U) \right) d\boldsymbol{M} \\
&\left. + \sum\nolimits_{\tau' \in \tau_{-f}} \boldsymbol{\lambda}(\tau') \int \frac{\boldsymbol{\beta}(\widetilde{\epsilon})}{N\overline{\boldsymbol{\beta}}} \left(v^{up}(\tau, \tau_{-f}, \epsilon^U + \widetilde{\epsilon}) - v^{up}(\tau, \tau_{-f}, \epsilon^U) \right) d\boldsymbol{M} \right\}, \\
&\rho v^{down}(\epsilon, n, \tau^D) \\
=& \max_{\widehat{\beta}} \left\{ \boldsymbol{\mu} \frac{\kappa}{2} \tau^D \epsilon v_z - \mathbf{w} \frac{1}{n} c_\beta(\widehat{\beta}) + \delta \sum\nolimits_{d \in D} [v^{down}\Big(\epsilon, n, \tau^D - \frac{1}{n} \sum\nolimits_{f' \in d} \tau_{f'}\Big) - v^{down}(\epsilon, n, \tau^D)] \right. \\
&+ \frac{\widehat{\beta}}{N\overline{\boldsymbol{\beta}}} \int \boldsymbol{\lambda}(\widetilde{\tau}) \left(\frac{\kappa}{2} \frac{\gamma}{n} \widetilde{\tau} \epsilon v_z + v^{down}\Big(\epsilon, n, \tau^D + \frac{\widetilde{\tau} + \sum_{f' \in \widetilde{j} \setminus \{\widetilde{f}\}} \tau_{f'}}{n}\Big) - v^{down}(\epsilon, n, \tau^D) \right) d\boldsymbol{M} \\
&\left. + \frac{(n-1)\boldsymbol{\beta}(\epsilon)}{N\boldsymbol{\beta}} \int \boldsymbol{\lambda}(\widetilde{\tau}) \left(v^{down}\Big(\epsilon, n, \tau^D + \frac{\widetilde{\tau} + \sum_{f' \in \widetilde{j} \setminus \{\widetilde{f}\}} \tau_{f'}}{n}\Big) - v^{down}(\epsilon, n, \tau^D) \right) d\boldsymbol{M} \right\}.
\end{aligned}$$

Note the state and value change after obtaining a new downstream industry in the expression for $v^{down}(\cdot)$. The term $\widetilde{\tau} + \sum_{f' \in \widetilde{j} \setminus \{\widetilde{f}\}} \tau_{f'}$ adds up all the output types of all firms in the new downstream industry. Only one firm (of $\widetilde{\tau}$) is the actual pathfinder that searches at $\boldsymbol{\lambda}(\widetilde{\tau})$, and the other firms' types are independent of $\widetilde{\tau}$. The $1/n$ adjustment is because of the definition of the state variable τ^D.

We aim to solve this system. The expression for $v^{in}(\cdot)$ is immediate, given as

$$v^{in}(\tau, \tau_{-f}, \epsilon) = \frac{\eta}{\rho} \left(\tau\epsilon \left(\boldsymbol{\mu}(1 - \frac{\kappa}{2}) v_z - \mathbf{w} c_\mu(\boldsymbol{\mu}) \right) + \boldsymbol{\mu} \frac{\kappa}{2} \frac{\sum_{\tau' \in \tau_{-f}} \tau' \epsilon}{n-1} v_z \right). \tag{B.3}$$

Then, we guess and verify $v^{up}(\tau, \tau_{-f}, \epsilon^U) = \tau\epsilon^U v_U + v_{\Delta U}(\tau, \tau_{-f})$. It is straightforward to show that v_U is the same as the benchmark case. The HJB for $v_{\Delta U}(\cdot)$ becomes

$$\rho v_{\Delta U}(\tau, \tau_{-f})$$

$$
\begin{aligned}
&= \max_{\widehat{\lambda}} \left\{ -\mathbf{w} c_\lambda(\widehat{\lambda}) + \widehat{\lambda} \int \frac{\boldsymbol{\beta}(\widetilde{\epsilon})}{N\overline{\boldsymbol{\beta}}} \left((1-\frac{\kappa}{2}) \frac{\gamma}{\widetilde{n}} \tau \widetilde{\epsilon} v_z + \tau \widetilde{\epsilon} v_U \right) d\boldsymbol{M} \right\} + \sum\nolimits_{\tau' \in \tau_{-f}} \boldsymbol{\lambda}(\tau') \int \frac{\boldsymbol{\beta}(\widetilde{\epsilon})}{N\overline{\boldsymbol{\beta}}} \tau \widetilde{\epsilon} v_U d\boldsymbol{M} \\
&= \max_{\widehat{\lambda}} \left\{ -\mathbf{w} c_\lambda(\widehat{\lambda}) + \widehat{\lambda} \int \frac{\boldsymbol{\beta}(\widetilde{\epsilon})}{N\overline{\boldsymbol{\beta}}} \tau \widetilde{\epsilon} \sum\nolimits_{\widetilde{n}} \left((1-\frac{\kappa}{2}) \gamma v_z + \widetilde{n} v_U \right) f_n(\widetilde{n}) dF_\epsilon(\widetilde{\epsilon}) \right\} \\
&\quad + \sum\nolimits_{\tau' \in \tau_{-f}} \boldsymbol{\lambda}(\tau') \int \frac{\boldsymbol{\beta}(\widetilde{\epsilon})}{N\overline{\boldsymbol{\beta}}} \tau \widetilde{\epsilon} v_U N dF_\epsilon(\widetilde{\epsilon}) \\
&= \max_{\widehat{\lambda}} \left\{ -\mathbf{w} c_\lambda(\widehat{\lambda}) + \widehat{\lambda} \tau \overline{\boldsymbol{\epsilon}_u} \left((1-\frac{\kappa}{2}) \frac{\gamma}{N} v_z + v_U \right) \right\} + \tau \overline{\boldsymbol{\epsilon}_u} v_U \sum\nolimits_{\tau' \in \tau_{-f}} \boldsymbol{\lambda}(\tau').
\end{aligned}
$$

The term $\tau \overline{\boldsymbol{\epsilon}_u} v_U \sum_{\tau' \in \tau_{-f}} \boldsymbol{\lambda}(\tau')$ is the expected flow value gain when others in j find a new upstream industry, and it strictly increases in the firm's own τ. The firm does not internalize the expected value $\widehat{\lambda}(n-1)\tau \overline{\boldsymbol{\epsilon}_u} v_U$ spilled over to others in j when choosing $\widehat{\lambda}$. The optimal search policy depends on output type only, such that

$$
\lambda^*(\tau) = (c'_\lambda)^{-1} \left(\tau \frac{\overline{\boldsymbol{\epsilon}_u}}{\mathbf{w}} \left((1-\frac{\kappa}{2}) \frac{\gamma}{N} v_z + v_U \right) \right). \tag{B.4}
$$

Equilibrium requires $\lambda^*(\tau) = \boldsymbol{\lambda}(\tau)$, so $v_{\Delta U}(\tau, \tau_{-f})$ satisfies

$$
v_{\Delta U}(\tau, \tau_{-f}) = \frac{1}{\rho} \left(\tau \overline{\boldsymbol{\epsilon}_u} \left(\boldsymbol{\lambda}(\tau) \frac{\gamma}{N} \left(1 - \frac{\kappa}{2} \right) v_z + \boldsymbol{\lambda}(\tau) v_U + \sum\nolimits_{\tau' \in \tau_{-f}} \boldsymbol{\lambda}(\tau') v_U \right) - \mathbf{w} c_\lambda(\boldsymbol{\lambda}(\tau)) \right) \tag{B.5}
$$

Lastly, we guess and verify $v^{down}(\epsilon, n, \tau^D) = \tau^D \epsilon v_D + v_{\Delta D}(\epsilon, n)$. It is again easy to see that v_D is the same as the benchmark case. The HJB for $v_{\Delta D}(\epsilon, n)$ becomes

$$
\begin{aligned}
\rho v_{\Delta D}(\epsilon, n) =& \max_{\widehat{\beta}} \left\{ -\mathbf{w} \frac{1}{n} c_\beta(\widehat{\beta}) + \frac{\widehat{\beta}}{N\overline{\boldsymbol{\beta}}} \int \boldsymbol{\lambda}(\widetilde{\tau}) \left(\frac{\kappa}{2} \frac{\gamma}{n} \widetilde{\tau} \epsilon v_z + \frac{\widetilde{\tau} + \sum_{f' \in \widetilde{j} \setminus \{\widetilde{f}\}} \tau_{f'}}{n} \epsilon v_D \right) d\boldsymbol{M} \right\} \\
&+ \frac{(n-1)\boldsymbol{\beta}(\epsilon)}{N\overline{\boldsymbol{\beta}}} \int \boldsymbol{\lambda}(\widetilde{\tau}) \frac{\widetilde{\tau} + \sum_{f' \in \widetilde{j} \setminus \{\widetilde{f}\}} \tau_{f'}}{n} \epsilon v_D d\boldsymbol{M} \\
=& \frac{1}{n} \max_{\widehat{\beta}} \left\{ -\mathbf{w} c_\beta(\widehat{\beta}) + \frac{\widehat{\beta}}{N\overline{\boldsymbol{\beta}}} \left(\int \boldsymbol{\lambda}(\widetilde{\tau}) \widetilde{\tau} \epsilon \left(\frac{\kappa}{2} \gamma v_z + v_D \right) N dF_\tau(\widetilde{\tau}) + \overline{\boldsymbol{\lambda}}(N-1) \overline{\tau} \epsilon v_D N \right) \right\} \\
&+ \frac{n-1}{n} \frac{\boldsymbol{\beta}(\epsilon)}{N\overline{\boldsymbol{\beta}}} \left(\int \boldsymbol{\lambda}(\widetilde{\tau}) \widetilde{\tau} \epsilon v_D N dF_\tau(\widetilde{\tau}) + (N-1) \overline{\tau} \overline{\boldsymbol{\lambda}} \epsilon v_D N \right) \\
=& \frac{1}{n} \max_{\widehat{\beta}} \left\{ -\mathbf{w} c_\beta(\widehat{\beta}) + \frac{\epsilon \widehat{\beta}}{\overline{\boldsymbol{\beta}}} \left(\overline{\boldsymbol{\tau}_d} \overline{\boldsymbol{\lambda}} \left(\frac{\kappa}{2} \gamma v_z + v_D \right) + (N-1) \overline{\tau} \overline{\boldsymbol{\lambda}} v_D \right) \right\} \\
&+ \frac{n-1}{n} \frac{\epsilon \boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}} \left(\overline{\boldsymbol{\tau}_d} \overline{\boldsymbol{\lambda}} + (N-1) \overline{\tau} \overline{\boldsymbol{\lambda}} \right) v_D.
\end{aligned}
$$

The negative congestion effect captured by $1/\overline{\boldsymbol{\beta}}$ is always present as in the benchmark model. Meanwhile, the term $\frac{n-1}{n} \frac{\epsilon \boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}} \left(\overline{\boldsymbol{\tau}_d} \overline{\boldsymbol{\lambda}} + (N-1) \overline{\tau} \overline{\boldsymbol{\lambda}} \right) v_D$ captures the positive

within-industry spillover effect, which is the expected flow value gain when others in j attracting a new downstream industry. The firm does not internalize the expected positive externality $\frac{n-1}{n}\frac{\epsilon\widehat{\beta}}{\overline{\boldsymbol{\beta}}}\left(\overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}+(N-1)\overline{\tau}\overline{\boldsymbol{\lambda}}\right)v_D$ when choosing $\widehat{\beta}$. The optimal attraction policy depends on input type only, such that

$$\beta^*(\epsilon)=(c'_\beta)^{-1}\left(\epsilon\frac{\overline{\boldsymbol{\lambda}}}{\overline{\boldsymbol{\beta}}\mathbf{w}}\left(\overline{\boldsymbol{\tau}_d}\frac{\kappa}{2}\gamma v_z+[\overline{\boldsymbol{\tau}_d}+(N-1)\overline{\tau}]v_D\right)\right). \tag{B.6}$$

Equilibrium requires $\beta^*(\epsilon)=\boldsymbol{\beta}(\epsilon)$, so $v_{\Delta D}(\epsilon,n)$ satisfies

$$v_{\Delta D}(\epsilon,n)=\frac{1}{\rho}\left(\frac{\epsilon\boldsymbol{\beta}(\epsilon)\overline{\boldsymbol{\lambda}}}{\overline{\boldsymbol{\beta}}}\left(\overline{\boldsymbol{\tau}_d}\frac{\kappa}{2}\frac{\gamma}{n}v_z+\left(\overline{\boldsymbol{\tau}_d}+(N-1)\overline{\tau}\right)v_D\right)-\mathbf{w}\frac{1}{n}c_\beta(\boldsymbol{\beta}(\epsilon))\right). \tag{B.7}$$

Combining all the pieces including Eqs. (B.3), (B.5) and (B.7) yields the complete expression for firm value achieved at Eqs. (B.2), (B.4) and (B.6):

$$V(\tau,\tau_{-f},\epsilon,z,\epsilon^U,\tau^D)=zv_z+v^{in}(\tau,\tau_{-f},\epsilon)+\tau\epsilon^U v_U+\tau^D\epsilon v_D+v_{\Delta U}(\tau,\tau_{-f})+v_{\Delta D}(\epsilon,n).$$

B.2.2 Growth and resource allocation in equilibrium

The aggregate flow of new links is deterministic, $\iint\boldsymbol{\lambda}(\omega)\frac{\boldsymbol{\beta}(\widetilde{\omega})}{\int\boldsymbol{\beta}(\omega')d\boldsymbol{M}}d\boldsymbol{M}d\boldsymbol{M}=\int\boldsymbol{\lambda}(\omega)d\boldsymbol{M}=N\overline{\boldsymbol{\lambda}}$, so the unweighted average degree satisfies $0=\dot{\boldsymbol{x}}=N\overline{\boldsymbol{\lambda}}-\delta\boldsymbol{x}$, i.e., $\boldsymbol{x}=\frac{N\overline{\boldsymbol{\lambda}}}{\delta}$.

Similar to the unweighted degrees, aggregation of match-quality-weighted links across nodes in either direction is equivalent to aggregation across links directly, so the average match-quality-weighted degree satisfies $\boldsymbol{X}=\frac{1}{N}\int\tau\epsilon^U d\boldsymbol{M}=\frac{1}{N}\int\tau^D\epsilon d\boldsymbol{M}$, where $\tau^D=\frac{1}{n}\sum_{d\in D}\sum f'\in d\tau_{f'}$ captures the fact that each firm is expected to contribute only $1/n$ of the industry-level input efficiency ϵ; the $1/N$ adjustment of aggregate firm measure ensures that $\boldsymbol{X}$ is the weighted average degree per firm. In the aggregate, $\boldsymbol{X}$ evolves deterministically as well, with outflow $\delta\boldsymbol{X}$ and inflow

$$\begin{aligned}
&\frac{1}{N}\iint\tau\left(\boldsymbol{\lambda}(\omega)+\sum\nolimits_{f'\in j\setminus\{\omega\}}\boldsymbol{\lambda}(\omega_{f'})\right)\frac{\boldsymbol{\beta}(\widetilde{\omega})}{\int\boldsymbol{\beta}(\omega'')d\boldsymbol{M}}\widetilde{\epsilon}\,d\boldsymbol{M}d\boldsymbol{M}\\
&=\left(\frac{1}{N}\int\tau\left(\boldsymbol{\lambda}(\omega)+\sum\nolimits_{f'\in j\setminus\{\omega\}}\boldsymbol{\lambda}(\omega_{f'})\right)d\boldsymbol{M}\right)\times\left(\int\frac{\widetilde{\epsilon}\boldsymbol{\beta}(\widetilde{\omega})}{\int\boldsymbol{\beta}(\omega'')d\boldsymbol{M}}d\boldsymbol{M}\right)\\
&=\left(\frac{1}{N}\int\tau\boldsymbol{\lambda}(\omega)d\boldsymbol{M}+\frac{1}{N}\overline{\tau}\int\sum\nolimits_{f'\in j\setminus\{\omega\}}\boldsymbol{\lambda}(\omega_{f'}))d\boldsymbol{M}\right)\times\left(\int\frac{\widetilde{\epsilon}\boldsymbol{\beta}(\widetilde{\omega})}{\int\boldsymbol{\beta}(\omega'')d\boldsymbol{M}}d\boldsymbol{M}\right)\\
&=\left(\frac{1}{N}\overline{\boldsymbol{\tau}_d}+\frac{N-1}{N}\overline{\tau}\right)N\overline{\boldsymbol{\lambda}}\overline{\boldsymbol{\epsilon}_u}.
\end{aligned}$$

At $\dot{\boldsymbol{X}} = 0$, we have $\boldsymbol{X} = \frac{1}{\delta} = \left(\frac{1}{N}\overline{\boldsymbol{\tau}_d} + \frac{N-1}{N}\overline{\tau}\right) N\overline{\boldsymbol{\lambda}}\overline{\boldsymbol{\epsilon}_u} = \left(\frac{1}{N}\overline{\boldsymbol{\tau}_d} + \frac{N-1}{N}\overline{\tau}\right)\boldsymbol{x}\overline{\boldsymbol{\epsilon}_u}$. The proof of $\overline{\boldsymbol{\tau}_d} > \overline{\tau}$ and $\overline{\boldsymbol{\epsilon}_u} > \bar{\epsilon}$ is the same as in the proof of Proposition 1.

Equilibrium requires that aggregate growth $\boldsymbol{g}$ and individual policies $\boldsymbol{\mu}, \boldsymbol{\lambda}(\cdot), \boldsymbol{\beta}(\cdot)$ are consistent as in Eq. (B.1). Therefore,

$$\begin{aligned}
\boldsymbol{g} =& \frac{1}{N}\left(\int \boldsymbol{\mu}\tau(\epsilon^U + \eta\epsilon)d\boldsymbol{M} + \iint \boldsymbol{\lambda}(\tau)\tau\frac{\boldsymbol{\beta}(\widetilde{\epsilon})\widetilde{\epsilon}}{\int \boldsymbol{\beta}(\epsilon')d\boldsymbol{M}}\frac{\gamma}{\widetilde{n}}d\boldsymbol{M}d\boldsymbol{M}\right) \\
=& \boldsymbol{\mu}\eta\overline{\tau\epsilon} + \frac{\boldsymbol{\mu}}{N}\int \tau\epsilon^U d\boldsymbol{M} + \frac{\gamma}{N} \times \left(\int \tau\boldsymbol{\lambda}(\tau)NdF_\tau(\tau)\right) \times \left(\int \frac{\widetilde{\epsilon}\boldsymbol{\beta}(\widetilde{\epsilon})}{N\overline{\boldsymbol{\beta}}}dF_\epsilon(\widetilde{\epsilon})\right) \times \left(\int \frac{\widetilde{n}}{\widetilde{n}}dF_n(\widetilde{n})\right) \\
=& \boldsymbol{\mu}\eta\overline{\tau\epsilon} + \boldsymbol{\mu}\boldsymbol{X} + \overline{\boldsymbol{\tau}_d}\frac{\gamma\overline{\boldsymbol{\lambda}}}{N}\overline{\boldsymbol{\epsilon}_u}.
\end{aligned}$$

Labor market clears at $R(\mathbf{w}) = \int \left(\tau(\epsilon^U + \eta\epsilon)c_\mu(\boldsymbol{\mu}) + c_\lambda(\boldsymbol{\lambda}(\tau)) + \frac{1}{n}c_\beta(\boldsymbol{\beta}(\epsilon))\right)d\boldsymbol{M} = N(\eta\overline{\tau\epsilon} + \boldsymbol{X})c_\mu(\boldsymbol{\mu}) + N\int c_\lambda(\boldsymbol{\lambda}(\tau))dF_\tau + \int c_\beta(\boldsymbol{\beta}(\epsilon))dF_\epsilon$.

B.2.3 Degree evolution and distributions

Consider an industry j of input type ϵ and with n firms of output types $(\tau_1, \ldots, \tau_n)$. Its outdegree x^U evolves as a birth-death process with birth rate $\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)$ and death rate δx^U; its indegree x^D evolves as a birth-death process with birth rate $\overline{\boldsymbol{\lambda}}n\boldsymbol{\beta}(\epsilon)/\overline{\boldsymbol{\beta}} \geq 0$ and death rate δx^D.

As in the benchmark case, we start from the outdegree or number of pathways found x^U, using the properties of birth-death processes. Its ergodic distribution has a probability mass function (pmf) that satisfies

$$\begin{aligned}
\boldsymbol{f}^U(x \mid \tau_1, \ldots, \tau_n) =& \frac{\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)}{\delta x}\boldsymbol{f}^U(x-1 \mid \tau_1, \ldots, \tau_n) = \left(\frac{\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)}{\delta}\right)^x \frac{1}{x!}\boldsymbol{f}^U(0 \mid \tau_1, \ldots, \tau_n), \\
\boldsymbol{f}^U(0 \mid \tau_1, \ldots, \tau_n) =& \left(\sum\nolimits_{x=0}^{\infty}\left(\frac{\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)}{\delta}\right)^x \frac{1}{x!}\right)^{-1} = e^{-\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)/\delta},
\end{aligned}$$

which is the pmf of a Poisson distribution with mean $\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)/\delta = \boldsymbol{x}\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)/(N\overline{\boldsymbol{\lambda}})$, so $x^U \mid \tau_1, \ldots, \tau_n \sim \text{Poisson}\left(\boldsymbol{x}\frac{\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)}{N\overline{\boldsymbol{\lambda}}}\right)$. The indegree or number of pathways received x^D is similar and omitted, such that $x^D \mid (\epsilon, n) \sim \text{Poisson}\left(\boldsymbol{x}\frac{n\boldsymbol{\beta}(\epsilon)}{N\overline{\boldsymbol{\beta}}}\right)$. Random search implies that, all else equal, an industry with more firms tend to have more upstream industries and more downstream ones.

We turn to each firm's relative productivity. Observe that firms in the same industry j share a common $\mathbb{E}[\epsilon^U \mid \tau_1, \ldots, \tau_n] = \boldsymbol{x}\frac{\sum_{f=1}^n \boldsymbol{\lambda}(\tau_f)}{N\overline{\boldsymbol{\lambda}}}\overline{\boldsymbol{\epsilon}_u}$ and a common $\mathbb{E}[\tau^D \mid$

$\epsilon, n] = \boldsymbol{x}\frac{n\boldsymbol{\beta}(\epsilon)}{N\overline{\boldsymbol{\beta}}}\frac{\overline{\boldsymbol{\tau}_d}+(N-1)\overline{\tau}}{n} = \boldsymbol{x}\frac{\boldsymbol{\beta}(\epsilon)}{\overline{\boldsymbol{\beta}}}\frac{\overline{\boldsymbol{\tau}_d}+(N-1)\overline{\tau}}{N}$. Consider a particular firm $f \in j$ with $\tau_f = \tau$. Its conditional expectation of match-quality-weighted supplier count (outdegree) becomes $\mathbb{E}[\tau\epsilon^U \mid \tau, \tau_{-f}] = \boldsymbol{x}\tau\frac{\boldsymbol{\lambda}(\tau)+\sum_{\tau'\in\tau_{-f}}\boldsymbol{\lambda}(\tau')}{N\overline{\boldsymbol{\lambda}}}\overline{\boldsymbol{\epsilon}_u}$. The expected law of motion of the firm's relative productivity at t is $dz = -\boldsymbol{g}zdt+\tau(\epsilon^U+\eta\epsilon)d\mathrm{P}(\boldsymbol{\mu},t)+\frac{1}{N}\gamma\overline{\boldsymbol{\epsilon}_u}d\mathrm{P}(\boldsymbol{\lambda}(\tau),t)$. It follows that, along a stationary equilibrium path, the long-run average of the firm's relative productivity is

$$\mathbb{E}[z \mid \tau, \tau_{-f}, \epsilon] = \frac{\boldsymbol{\mu}\eta\tau\epsilon + \boldsymbol{\mu}\mathbb{E}[\tau\epsilon^U \mid \tau, \tau_{-f}] + \tau\boldsymbol{\lambda}(\tau)\frac{\gamma}{N}\overline{\boldsymbol{\epsilon}_u}}{\boldsymbol{g}} = \frac{\boldsymbol{\mu}\eta\tau\epsilon + \boldsymbol{\mu}\boldsymbol{x}\tau\frac{\boldsymbol{\lambda}(\tau)+\sum_{\tau'\in\tau_{-f}}\boldsymbol{\lambda}(\tau')}{N\overline{\boldsymbol{\lambda}}}\overline{\boldsymbol{\epsilon}_u} + \tau\boldsymbol{\lambda}(\tau)\frac{\gamma}{N}\overline{\boldsymbol{\epsilon}_u}}{\boldsymbol{\mu}\eta\overline{\tau\epsilon} + \boldsymbol{\mu}\boldsymbol{X} + \overline{\boldsymbol{\tau}_d}\frac{\gamma}{N}\overline{\boldsymbol{\epsilon}_u}},$$

which is strictly increasing in ϵ when $\eta > 0$, strictly increasing and convex in τ, strictly increasing in each peer firm's $\tau' \in \tau_{-f}$, and is super-modular in τ and τ'. In the case of $\gamma = 0$ and $\eta = 0$, the expression simplifies to $\mathbb{E}[z \mid \tau, \tau_{-f}, \epsilon] = \frac{\tau\cdot\left(\boldsymbol{\lambda}(\tau)+\sum_{\tau'\in\tau_{-f}}\boldsymbol{\lambda}(\tau')\right)}{\left(\frac{1}{N}\overline{\boldsymbol{\tau}_d}+\frac{N-1}{N}\overline{\tau}\right)\cdot N\overline{\boldsymbol{\lambda}}}$.

B.3 Other proofs

Proof of Lemma B.1. Given any value function $V(\cdot)$, the generic form of joint innovation surplus is $S_\mu(\omega, (\omega_{uh})_{uh=1}^{x^U+1}) = V(\omega') - V(\omega) + \sum_{uh=1}^{x^U+1}\big(V(\omega'_{uh}) - V(\omega_{uh})\big)$, where $(\omega_{uh})_{uh=1}^{x^U+1}$ represents the list of matched suppliers in each of the x^U upstream industries and own industry, notation $'$ indicates firm states after the innovation. For the innovator ω, the only change is the relative productivity jumps from z to $z+\tau(\epsilon^U+\eta\epsilon)$. For each matched supplier ω_{uh}, its state remains unchanged. Therefore, with $V(\omega) = zv_z + v(\omega_{-z})$, the joint surplus to be split simplifies to $S_\mu(\tau, \epsilon^U + \eta\epsilon) = \tau(\epsilon^U + \eta\epsilon)v_z$. Recall that the innovator ω has the within-industry replacement option to use an unmatched supplier instead of the matched one at the proportional match quality cost of κ. Formally, the coalitional game shall be defined as (I, φ), where I is the set of $x^U + 2$ players including the innovator ω and all matched suppliers, and $\varphi : \mathcal{P}(I) \to \mathbb{R}_+$ is the characteristic function specifying the payoff to each possible coalition $\widetilde{I} \subseteq I$. Specifically, $\forall\widetilde{I} \subseteq I$, if $\omega \notin \widetilde{I}$, i.e., without the innovator, then $\varphi(\widetilde{I}) = 0$; otherwise, whenever $\omega \in \widetilde{I}$, $\varphi(\widetilde{I}) = \tau(\epsilon^U + \eta\epsilon)v_z - \tau\sum_{\omega_{uh}\in I\setminus\widetilde{I}}\kappa\epsilon_u v_z,$, where $\epsilon_u = \eta\epsilon$ when the supplier of own-industry knowledge is not in $\widetilde{I}$. It means that inputs in $\widetilde{I}$ are used at full efficiency as they are provided by the matched suppliers, and the innovator has access to unmatched suppliers for inputs in $I \setminus \widetilde{I}$ at the cost κ.

Now we turn to each firm's marginal contribution. Clearly, the innovator ω's

marginal contribution to any $\widetilde{I} \not\ni \omega$ is always $\Delta(\widetilde{I}, \omega) = \varphi(\widetilde{I}) - 0 = \varphi(\widetilde{I})$, and each matched supplier ω_{uh}'s marginal contribution to any $\widetilde{I} \not\ni \omega_{uh}$ is $\Delta(\widetilde{I}, \omega_{uh}) = 0$ if $\omega \notin \widetilde{I}$, and $= \varphi(\widetilde{I} \cup \{\omega_{uh}\}) - \varphi(\widetilde{I}) = \kappa\tau\epsilon_u v_z$ if $\omega \in \widetilde{I}$. The Shapley value solution $Sh(i; I, \varphi)$ for each player $i \in I$ is formally defined by its average marginal contribution to any coalition, $Sh(i; I, \varphi) = \frac{1}{|I|!}\sum_{R\in\mathcal{R}(I)}\Delta(\widetilde{I}_i(R), i)$, where $|I|$ is the number of players, R is a permutation or ordering of players, $\mathcal{R}(I)$ is the set of all $|I|!$ possible orderings of I, and $\widetilde{I}_i(R)$ is the set of players before i in ordering R. Observe that the assumption of separable input contribution makes the exact ordering of all I irrelevant; only the relative positions of the innovator ω and a particular supplier ω_{uh} matters: ω appears before ω_{uh} in half of all orderings and after in the other half. Therefore, $Sh(\omega_{uh}; I, \varphi) = S^U(\tau, \epsilon_u) = \frac{\kappa}{2}\tau\epsilon_u v_z$, and $Sh(\omega; I, \varphi) = S_\mu(\tau, \epsilon^U + \eta\epsilon) - \sum_{u\in U} S^U(\tau, \epsilon_u) - S^U(\tau, \eta\epsilon) = \left(1 - \frac{\kappa}{2}\right)\tau(\epsilon^U + \eta\epsilon)v_z$.

Bilateral bargaining is similar. The joint surplus when ω finds $\widetilde{\omega}$ is $S_\lambda(\omega, \widetilde{\omega}) = V(\omega') - V(\omega) + V(\widetilde{\omega}') - V(\widetilde{\omega}) = \frac{\gamma}{\widetilde{n}}\tau\widetilde{\epsilon}v_z + v(\omega'_{-z}) - v(\omega_{-z}) + v(\widetilde{\omega}'_{-z}) - v(\widetilde{\omega}_{-z})$, where the additional state changes are at the industry level, such that the pathfinder ω's upstream set becomes $U \cup \{\widetilde{j}\}$ and in the receiver $\widetilde{\omega}$'s downstream set $\widetilde{D} \cup \{j\}$, where j and $\widetilde{j}$ represent the respective industries of ω and $\widetilde{\omega}$. Whether the finder ω turns to another firm as the receiver in $\widetilde{\omega}$'s industry $\widetilde{j}$ does not affect these states. Therefore, the additional value of sticking to $\widetilde{\omega}$ is simply the switching cost saved, i.e., $\kappa\frac{\gamma}{\widetilde{n}}\tau\widetilde{\epsilon}v_z$. Consequently, given symmetric bargaining weights, the receiver $\widetilde{\omega}$ gets from the finder ω a transfer of $S^U(\tau, \tau\widetilde{\epsilon}/\widetilde{n}) = \frac{\kappa}{2}\frac{\gamma}{\widetilde{n}}\tau\widetilde{\epsilon}v_z = \arg\max_{\widehat{T}}(\kappa\frac{\gamma}{\widetilde{n}}\tau\widetilde{\epsilon}v_z - \widehat{T})(\widehat{T})$. □

Proof of Proposition 5. See Appendix B.2.2. □

C Identification and Quantitative Analysis

C.1 Proof of Corollary 2

Proof. It follows the same logic as in the proof of Proposition 3 that if $c_\beta(\widehat{\beta}) \propto \widehat{\beta}^{1+\chi_\beta}$, $\chi_\beta > 0$, then $\overline{\boldsymbol{\epsilon}_u} = \frac{\int \epsilon^{1+1/\chi_\beta} dF_\epsilon}{\int \epsilon^{1/\chi_\beta} dF_\epsilon}$. Consequently, if $\epsilon \sim \text{Exp}(\theta)$ such that $F_\epsilon(\epsilon) = 1 - e^{-\theta\epsilon}$, then $\bar{\epsilon} = 1/\theta$, and $\overline{\boldsymbol{\epsilon}_u} = \frac{\int \epsilon^{1+1/\chi_\beta}\theta e^{-\theta\epsilon} d\epsilon}{\int \epsilon^{1/\chi_\beta}\theta e^{-\theta\epsilon} d\epsilon} = \frac{\theta^{-(1+1/\chi_\beta)}\Gamma(2+1/\chi_\beta)}{\theta^{-1/\chi_\beta}\Gamma(1+1/\chi_\beta)} = \bar{\epsilon}(1 + 1/\chi_\beta)$, where $\Gamma(\cdot)$ is the gamma function. □

C.2 Step-by-step identification

We copy the moment list here in Table C.1 here for easy reference. Assuming availability of "ideal data" and that the model at its stationary equilibrium is the true data-generating process (DGP), we show step-by-step identification in what follows.

We caution that we do *not* follow the identification procedure when calibrating the model, as the model is intentionally stylized and unlikely to be the true DGP. Instead, we use the standard GMM. Moreover, we are aware of the discrepancy between a continuous-time model (jointly governed by several endogenous Poisson processes in equilibrium) and discrete data observations. A more rigorous approach requires to match the individual-level state-transition probabilities induced by these endogenous Poisson processes given the unit time length. See, for example, Arcidiacono et al. (2016) and Li et al. (2024) in less complex settings. In this model, we take the linear-approximation shortcut and treat the Poisson rates (e.g., $\boldsymbol{\mu}$, $\boldsymbol{\lambda}(\tau)$, δ) simply as probabilities per time unit. The main reason is that we do not have sufficient data to estimate the transition probabilities given the high-dimensional firm states.

Identification.

Step 1. We read from the table that $\frac{\mathbf{w}R(\mathbf{w})}{100} = \frac{\mathbf{w}^{1+\sigma}}{100\varsigma^{\sigma}} = m(1) \times m(6) = \frac{1}{\nu} - m(3)$, which immediately identifies $\nu = 1/(m(3) + m(1) \times m(6))$, so L is also known, and $\boldsymbol{\pi} = \frac{100}{N\nu}$. We also have $m(2) = \frac{\nu}{\nu-1}\frac{\mathbf{w}^{\sigma}}{100\varsigma^{\sigma}}$. It follows that $\mathbf{w} = \frac{m(1)\times m(6)}{m(2)}\frac{\nu-1}{\nu}$ and $R(\mathbf{w}) = 100\frac{m(1)\times m(6)}{\mathbf{w}}$. With a predetermined σ, we obtain ς.

Step 2. Given $\boldsymbol{\pi}$ and $m(4) = \boldsymbol{g}$, we have v_z. Then, using $m(6)$ again yields κ.

Step 3. With $c_\mu(\widehat{\mu}) = \frac{c_\mu}{2}\widehat{\mu}^2$, $m(7) = \boldsymbol{\mu}$ identifies c_μ because $\boldsymbol{\mu} = \frac{1}{c_\mu}\frac{v_z}{\mathbf{w}}(1 - \frac{\kappa}{2})$.

Step 4. Using $m(8) = \boldsymbol{x} = \frac{N}{\delta}\overline{\boldsymbol{\lambda}} = \frac{N}{\delta}m(7)$, we have δ. Then, v_U and v_D follow.

Step 5. The four remaining parameters are $(\overline{\tau} = \bar{\epsilon}, c_\lambda, \chi_\lambda, \gamma)$, jointly identified by a system of four equations Eqs. (C.1) to (C.4)below.

(i) Observe that $\frac{\overline{\epsilon_u}}{\bar{\epsilon}} = 1 + 1/\chi_\beta$ is already known. Meanwhile, any random variable from an exponential distribution has $\overline{\tau^k} := \mathbb{E}[\tau^k] = \Gamma(k+1)\overline{\tau}^k$, $\forall k > 0$, or $\overline{\tau^k}/\overline{\tau}^k = \Gamma(k+1)$, where $\Gamma(\cdot)$ is the gamma function. Equilibrium search policy $\boldsymbol{\lambda}(\tau)$ satisfying Eq. (B.4) becomes $\boldsymbol{\lambda}(\tau) = \frac{1}{c_\lambda}\left(\frac{\tau\overline{\epsilon_u}}{\mathbf{w}}\left(\frac{\gamma}{N}\left(1 - \frac{\kappa}{2}\right)v_z + v_U\right)\right)^{1/\chi_\lambda} =$

Table C.1: Targeted moments

m	description	expression
(1)	R&D expenditure to GDP ratio	$(\mathbf{w}R(\mathbf{w}) + \frac{\kappa}{2}N\boldsymbol{g}v_z)/100$
(2)	researcher employment share	$R(\mathbf{w})/L$
(3)	profits to GDP ratio	$\Pi/100 = \frac{1}{\nu} - \mathbf{w}R(\mathbf{w})/100$
(4)	innovation rate	$\boldsymbol{g}$
(5)	firms' average innovation frequency	$\boldsymbol{\mu}$
(6)	researcher compensation share of total R&D	$\mathbf{w}R(\mathbf{w})/(\mathbf{w}R(\mathbf{w}) + \frac{\kappa}{2}N\boldsymbol{g}v_z)$
(7)	firms' average probability of finding a pathway	$\overline{\boldsymbol{\lambda}}$
(8)	avg. degree of industry-level knowledge network	$\boldsymbol{x}$
(9)	avg. pathfinder and receiver's sales increase at pathway formation	$\frac{\gamma}{N}\overline{\tau}(1+1/\chi_\lambda)\overline{\epsilon}(1+1/\chi_\beta)$
empirical counterparts for externally calibrated parameters		
	avg. number of firms in an industry	N
	avg. ratio: own-industry ($j \to j$) citations / avg. citations per any other cited industry ($j \to j' \neq j$)	η
	avg. citations received per input patent / avg. citations	$\overline{\boldsymbol{\epsilon}_u}/\overline{\epsilon} \implies \chi_\beta = (\overline{\boldsymbol{\epsilon}_u}/\overline{\epsilon} - 1)^{-1}$

Normalization: $\boldsymbol{P}(t)\boldsymbol{C}(t) = 100$, $L = 100\frac{\nu-1}{\nu}$.

$\frac{(\tau\bar{\epsilon})^{1/\chi_\lambda}}{c_\lambda}\left(\frac{1+1/\chi_\beta}{\mathbf{w}}\left(\frac{\gamma}{N}\left(1-\frac{\kappa}{2}\right)v_z + v_U\right)\right)^{1/\chi_\lambda}$, which implies

$$m(7) = \overline{\boldsymbol{\lambda}} = \frac{\Gamma(1+1/\chi_\lambda)(\overline{\tau\epsilon})^{1/\chi_\lambda}}{c_\lambda}\left(\frac{1+1/\chi_\beta}{\mathbf{w}}\left(\frac{\gamma}{N}\left(1-\frac{\kappa}{2}\right)v_z + v_U\right)\right)^{1/\chi_\lambda}. \quad \text{(C.1)}$$

Notice that $\boldsymbol{\lambda}(\tau)$ can be written as $\boldsymbol{\lambda}(\tau) = \lambda\tau^{1/\chi_\lambda}$ with λ being the endogenous coefficient such that $\lambda^{\chi_\lambda} = \frac{\bar{\epsilon}(1+1/\chi_\beta)}{\mathbf{w}c_\lambda^{\chi_\lambda}}\left(\frac{\gamma}{N}\left(1-\frac{\kappa}{2}\right)v_z + v_U\right)$, and we have $\frac{\overline{\boldsymbol{\tau}_d}}{\overline{\tau}} = \frac{\overline{\tau^{1+1/\chi_\lambda}}}{\overline{\tau}\cdot\overline{\tau^{1/\chi_\lambda}}} = 1+1/\chi_\lambda$. Similarly, with $c_\beta(\widehat{\beta}) = \frac{c_\beta}{1+\chi_\beta}\widehat{\beta}^{1+\chi_\beta}$, equilibrium attraction policy in Eq. (B.6) becomes $\boldsymbol{\beta}(\epsilon) = \beta\epsilon^{1/\chi_\beta} = \left(\frac{1}{c_\beta\mathbf{w}\beta\overline{\epsilon^{1/\chi_\beta}}}\left(\overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}\gamma\frac{\kappa}{2}v_z + \overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}v_D + \overline{\tau}\overline{\boldsymbol{\lambda}}(N-1)v_D\right)\right)^{1/\chi_\beta}\epsilon^{1/\chi_\beta}$, which implies $\beta^{1+\chi_\beta} = \frac{1}{c_\beta\mathbf{w}\overline{\epsilon^{1/\chi_\beta}}}\left(\overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}\gamma\frac{\kappa}{2}v_z + \overline{\boldsymbol{\tau}_d}\overline{\boldsymbol{\lambda}}v_D + \overline{\tau}\overline{\boldsymbol{\lambda}}(N-1)v_D\right)$, where c_β cannot be separately identified from β without observations of $\boldsymbol{\beta}(\epsilon)$. Normalizing $c_\beta = 1$, we have $\beta^{1+\chi_\beta} = \frac{\overline{\tau}\overline{\boldsymbol{\lambda}}}{\mathbf{w}\bar{\epsilon}^{1/\chi_\beta}\Gamma(1+1/\chi_\beta)}\left((1+1/\chi_\lambda)\gamma\frac{\kappa}{2}v_z + (1+1/\chi_\lambda)v_D + (N-1)v_D\right)$.

(ii) Equilibrium growth rate in Eq. (20) can be written as

$$m(4) = \boldsymbol{g} = \overline{\tau\epsilon}\left(\boldsymbol{\mu}\eta + \boldsymbol{\mu}\left(\frac{1}{N\chi_\lambda} + 1\right)\boldsymbol{x}(1+1/\chi_\beta) + \overline{\boldsymbol{\lambda}}(1+1/\chi_\lambda)\frac{\gamma}{N}(1+1/\chi_\beta)\right). \quad \text{(C.2)}$$

(iii) The market clearing condition for research labor becomes

$$R(\mathbf{w}) = N\overline{\tau\epsilon}\left(\eta + \left(\frac{1}{N\chi_\lambda} + 1\right)\boldsymbol{x}(1+1/\chi_\beta)\right)\frac{c_\mu}{2}\boldsymbol{\mu}^2 + N\frac{c_\lambda^{\chi_\lambda}\lambda^{1+\chi_\lambda}}{1+\chi_\lambda}\overline{\tau^{1+1/\chi_\lambda}} + \frac{\beta^{1+\chi_\beta}}{1+\chi_\beta}\overline{\epsilon^{1+1/\chi_\beta}} \tag{C.3}$$

where λ and β are functions of c_λ, χ_λ and $\overline{\tau} = \overline{\epsilon}$.

(iv) Construct $m(9)$ as

$$\begin{aligned} m(9) =& (\text{avg. \% jump in pathfinder's and receiver's sales at the formation of a pathway}) \\ & \times (\text{avg. finder's relative size} + \text{avg. receiver's relative size}), \end{aligned}$$

where relative sizes are measured as the finder's and the receiver's sizes in sales relative to an average firm's sales. We turn to the model. Let z_f and z_r be the relative productivities of a pathfinder and a receiver. Prior to the new pathway, their average flow profit is $\boldsymbol{\pi}(z_f + z_r)/2$. Immediately after the new pathway, their jumps in flow profits are $\gamma\tau_f\frac{\epsilon_r}{n_r}\left(1 - \frac{\kappa}{2}\right)\boldsymbol{\pi}$ and $\gamma\tau_f\frac{\epsilon_r}{n_r}\frac{\kappa}{2}\boldsymbol{\pi}$, where we interpret the receiver's surplus as its value gain due to productivity jump. It follows that their average jump in flow profits is $\boldsymbol{\pi}\gamma\tau_f\frac{\epsilon_r}{n_r}/2$, which can be written as a fraction of their average flow profit $\gamma\tau_f\frac{\epsilon_r}{n_r}/(z_f + z_r)$, which has $m(9)$ as the empirical counterpart. We make some observations. First, recall the definition of relative productivity such that $\overline{z} = 1$ across all firms, so $z_f + z_r$ is finder and receiver's sizes relative to an average firm measured in profits or sales. Second, the average of all receivers' ϵ_r should be $\overline{\boldsymbol{\epsilon}_u} = \overline{\epsilon}(1 + 1/\chi_\beta)$. Similarly, the average of all finders' τ_f should be $\overline{\tau}(1 + 1/\chi_\lambda)$, as discussed in the previous identification step. It follows that we can obtain γ as

$$m(9) = \frac{\gamma}{N} \times \overline{\tau}(1 + 1/\chi_\lambda)\overline{\epsilon}(1 + 1/\chi_\beta). \tag{C.4}$$

C.3 Computation of the stationary equilibrium

The steps to solve for the equilibrium in Appendix B already suggest the computation procedure. In practice, solving for the stationary equilibrium boils down to find the solution to a simple fixed-point problem: $(\mathbf{w}, \boldsymbol{g}) = \mathcal{T}(\mathbf{w}, \boldsymbol{g})$, where $\mathcal{T}$ is a mapping on $[0, w_{\max}] \times [0, g_{\max}]$ with sufficiently large but finite $w_{\max}, g_{\max} > 0$. The mapping $\mathcal{T}$ is defined to follow Appendix B.2 closely. Given a pair of (w, g), find the value-function coefficients $v_z(w, g), v_U(w, g), v_D(w, g)$ and the optimal innovation

policy $\mu^*(w,g)$ as in Eq. (B.2). Proceed to solve for $\lambda^*(\tau;w,g)$ and $\beta^*(\epsilon;w,g)$ simultaneously as in Eqs. (B.4) and (B.6). Leveraging the functional-form assumptions, this is straightforward. Specifically, $\overline{\boldsymbol{\epsilon}_u}$ in equilibrium can be solved outside of the equilibrium, as established by Corollary 2. Then we get $\lambda^*(\tau;w,g) = \lambda(w,g)\tau^{1/\chi_\lambda}$, and $\beta^*(\epsilon;w,g) = \beta(w,g)\epsilon^{1/\chi_\beta}$ follows. At this stage, we have each component of growth rate in Eq. (20) updated using Eq. (21), denoted as g'. We also obtain the aggregate demand for research labor $R_d(w,g)$, which corresponds to an updated wage denoted as w'. Thus, $\mathcal{T}(w,g) = (w',g')$, with the rest of the equilibrium solved in the interim. Observe that $\mathcal{T}$ is continuous, so the Brouwer fixed-point theorem guarantees the existence of a stationary equilibrium. A numerical solution is $(\widehat{\mathbf{w}},\widehat{\boldsymbol{g}}) \in \arg\min_{w,g} \|(w,g) - \mathcal{T}(w,g)\|_2$ with $\|\cdot\|_2$ being the Euclidian norm.

D Additional Figures

D.1 More on Facts 3 and 4

Fig. D.1.1 is an alternative to Fig. 3 and considers the citation weighted patent stock. It corresponds to Fact 3. Figs. D.1.2 to D.1.6 correspond to the other measures of firm performance mentioned in Fact 4, including sales, employment, market value, capital stock, and TFP.

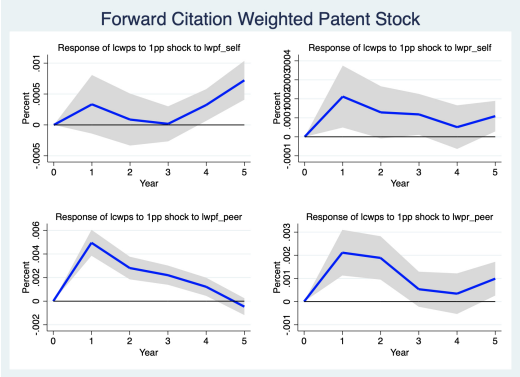


Figure D.1.1: Knowledge Pathway Creation and Firms' Citation Weighted Patent Stock

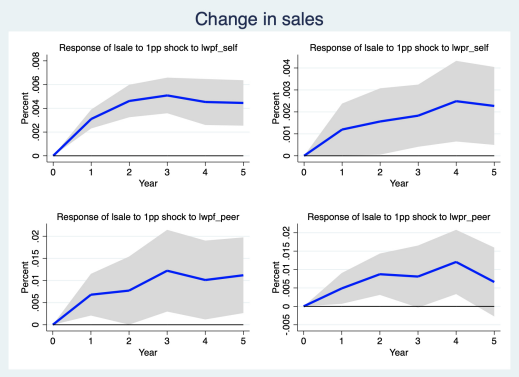


Figure D.1.2: Knowledge Path Creation and Firm Sales

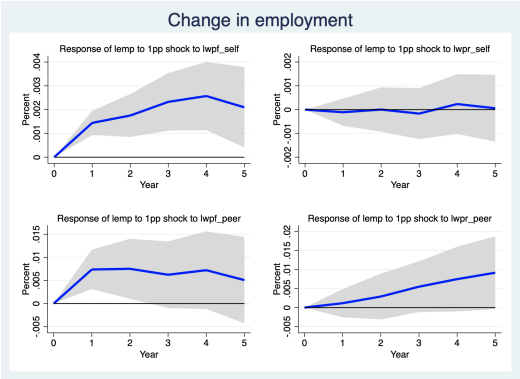


Figure D.1.3: Knowledge Path Creation and Firm Employment

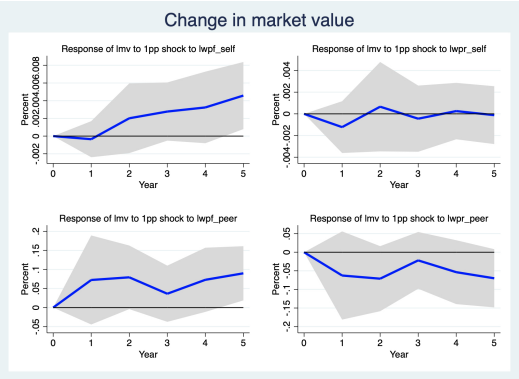


Figure D.1.4: Knowledge Path Creation and Firm Market Value

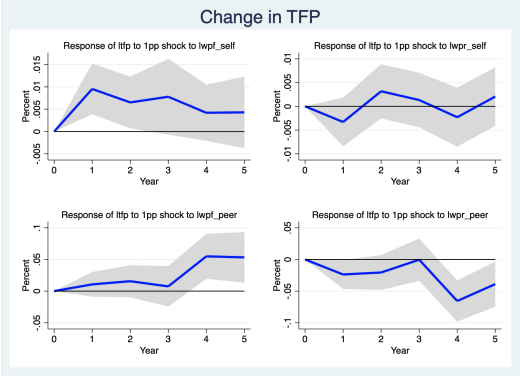


Figure D.1.5: Knowledge Path Creation and Firm TFP

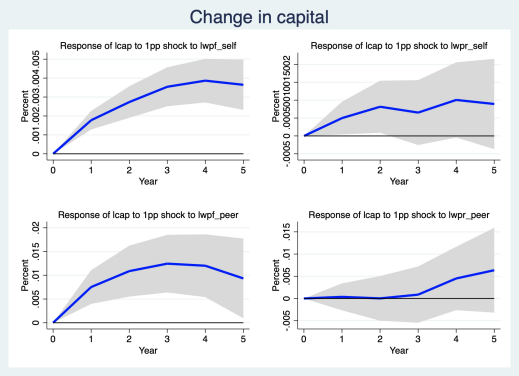


Figure D.1.6: Knowledge Path Creation and Firm Capital Stock

D.2 A subset of robustness checks

In the regressions using Compustat public firms, industry (SIC4), firm and year fixed effects are included. In the regressions using patenting firms, firm and year fixed effects are included. Standard errors are clustered at SIC4 and year level. lwpf/lwpr are log scaled patent stock weighted number of new outward/inward citations from/to firm's technology space. Self/peers means new citation links made by self/other firms. Lagged log-scaled employment, capital and the dependent variable are included in the regression, but not reported.

To save space, we only show the counterparts to Figs. 3 and 4 using different measures of knowledge pathway creation from one patent category to another: the first 10 citation links (Figs. D.2.1 and D.2.2), the first 100 citation links (Figs. D.2.3

and D.2.4), and the first citation link after at least a decade of no citations (Figs. D.2.5 and D.2.6). Other results are available upon request.

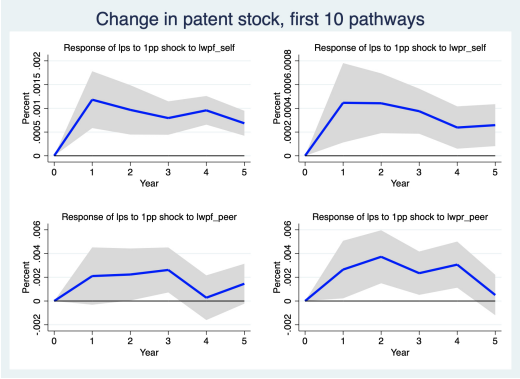


Figure D.2.1: First 10 Citations through a Knowledge Pathway and Firm Patent Stock

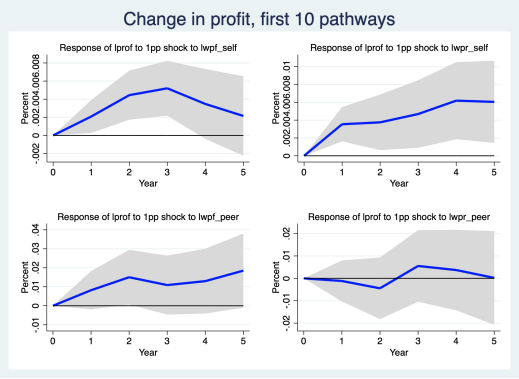


Figure D.2.2: First 10 Citations through a Knowledge Pathway and Firm Profit

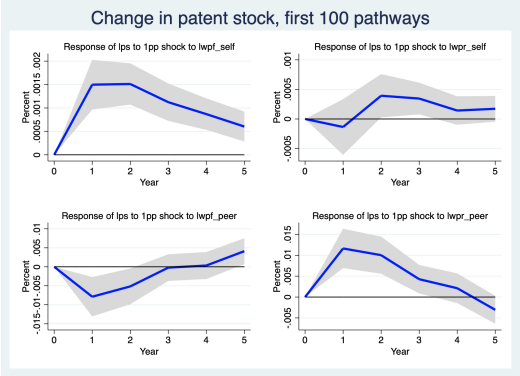


Figure D.2.3: First 100 Citations through a Knowledge Pathway and Firm Patent Stock

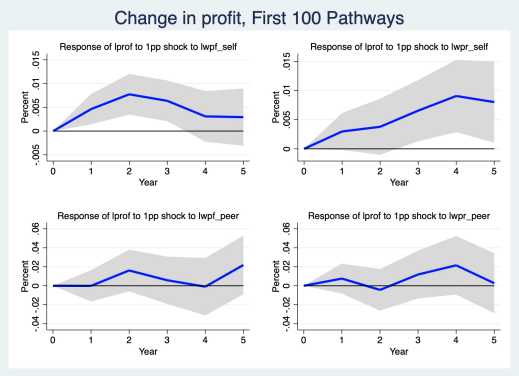


Figure D.2.4: First 100 Citations through a Knowledge Pathway and Firm Profit

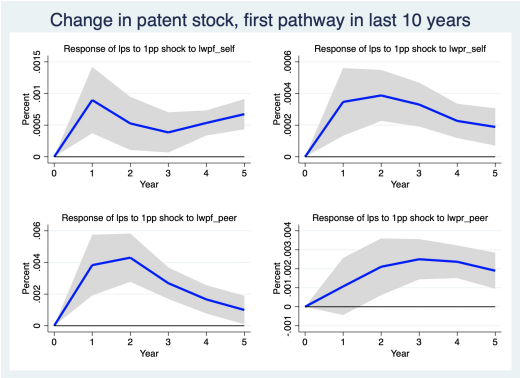


Figure D.2.5: Knowledge New Pathway In the Last 10 years and Firm Patent Stock

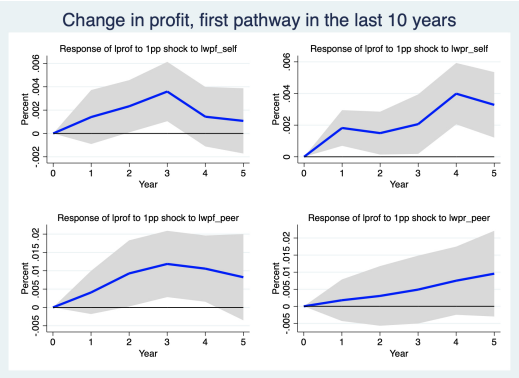


Figure D.2.6: First Pathway in the Last 10 Years and Firm Profit