# Dynamic Formation of Knowledge Networks and Innovating Firm

Jie Cai and Can Tian*

February 4, 2021

**Abstract**

We document a series of new facts about the very first firms and patents that form new edges in the directed citation networks across patent categories. We call them pathfinder firms and patents. First, the typical pathfinder firms are very larger firms. Second, the average pathfinder patents have higher quality than other patents in different quality measures. Third, firms innovate faster, market value, profit and productivity increase in the future, when they invent a large number of pathfinder patents or are cited by pathfinder patents currently. Fourth, new citation links generate positive externalities to peer firms that innovate in nearby technology space. We then build a dynamic formation model of knowledge network, where new citations form through both quality based preferential attachment, exact and mutated copying of parent patent's citations. The model rationalizes the firm level empirical facts. After calibrated to the patent citation data, the model sheds lights on the causes of productivity show down: rising number of citation per patent, marginal cost of patent quality and research wage have driven up the cost to innovate since 1976. Effective R&D policies need to decrease these cost factors.

JEL Classification codes: O31 O33 O41 D85

Key words: innovation, knowledge networks, patent citations, dynamic network formation, R&D policy

**Very preliminary, please do not circulate

## 1 Introduction

Innovating firms build the cross sector knowledge networks, they also rely on this networks to create, absorb knowledge and expand into new areas of the technology space. How do firms build the knowledge networks? What kind of firms are more likely to build cross sector new links in the knowledge networks? Do firms benefit from the new knowledge links built by themselves and other firms? What kind of dynamic formation process of the knowledge networks can generate

*Cai: School of Economics, Shanghai University of Finance and Economics, april.cai@gmail.com; Tian: Department of Economics, University of North Carolina at Chapel Hill, tiancan@email.unc.edu.

the topology of the real knowledge networks and the cross sector knowledge flows observed in the patent citation data?

To answer the above questions, we focus on a special type of patents, pathfinder patents, and their implication to patenting firms' performances. Patent A in sector i that cites a patent B in sector j is called a pathfinder patent, if A is the first patent in sector i that has ever cited a sector j patent. We call patent B a path-receiver patent. The owner firm of patent A (B) is called pathfinder (path-receiver) firm. One example of pathfinder patent is US Patent NO. 9,116,285 filed by 3M Innovative Properties Company in 2015, this light directing film is used in applications such as TV, computer and cell phones displays to increase brightness and reduce the overall thickness. This patent is the first in class 362 (Illumination) that cited another patent NO. 7,140,812 in class 407 (Cutter). Patent 7,140,812 owned by the same firm is a Diamond cutting tool with a multi-tipped diamond. The diamond cutting tool in Figure 2 is used to cut the surface of the light directing film shown in Figure 3, this design simplifies and improves the creation of a micro-replication tool using a diamond.

The cross-category citation networks are built step by step by such pathfinder patents and firms over time. Studying pathfinder patents help us understand the dynamic formation process of knowledge networks and the implications of knowledge network dynamics on innovating firms. Changes to knowledge network structure affect firms who innovation at neighbourhood location in the technology space, because they absorb innovation input and receive royalty payment through knowledge linkages. In the previous example, not only 3M benefit from this new citation path from class 362 to class 407. On one hand, peer firms in class 362 will learn to use diamond cutting tool in the fabrication of light directing films, hence make their light directing films brighter and thinner in the future; On the other hand, peer firms in class 407 expect that their knowledge on other diamond cutting tools has potential to be applied in the manufacturing of light directing films. Therefore other class 407 patents become more valuable.

In this paper we first present empirical facts on the the real consequences of knowledge networks dynamics using matched Compustat and USPTO data. Then we construct a firm innovation model, where new citations are made by both quality attracted preferential attachment and exact or mutated copy of parental citations, to explain the empirical firm level facts. Lastly, the calibrated model helps us diagnose the reasons for productivity slow down and design effective R&D policy.

In the empirical section, we document several facts about pathfinder patents and their real impact on innovating firms. First, Pathfinder patents have higher quality in different measures than average patents. They receive 57% more forward citations, score 52% higher in originality and 16% higher in generality, span 0.16 more CPC patent classes, cite 1.9 year younger patents, cite 27% more scientific literature than average patents. However, pathfinder patents score lower than average in backward-citations' pedigree, grant lag and novel word count, because pathfinder patents' novelty comes from combining a new set of knowledge inputs together to make an existing product better, instead of citing other popular patents and proposing a brand new idea. Path-receiver patents have similar or even higher quality measures in all dimensions as pathfinder patents, except for a fewer backward citations to scientific literature. That is because inventors prefer well established and simple patents that are far away from research frontier, when they use knowledge input from a new sector for the first time. Second, pathfinder firms are very large firms in the patent dataset. On average pathfinder firms own 251 patent stocks and have patented in 6 technology categories; while average firms in the patent data own 46 patents and have patented in 1.84 technology categories. Path-receiver firms are even larger than pathfinder firms, on average they own 415 patents and have patented in 8 technology categories.

Third, a patenting firm's innovate rate is higher in the future 5 years, when there are more new cross-category inward and outward citation links made on this firm's technology space either by itself or peer firms in current year. The new cross-category citation links mainly help firms expand into new technology categories, rather than grow intensive within existing categories. Fourth, in linked CRSP/Compustat data, we find that public listed firm's real performances, such as market value, employment, capital, sales, profit and TFP, increase in the coming 5 years, when there are more new cross-category inward and outward links made on the firm's technology space either by itself or peer firms in current financial year. Fifth, we find that citation counts across technology categories, like goods flow across countries, follow a gravity model, which is positively related to the patent stocks in both categories. Sixth, firm growth becomes more volatile, when the firm's innovation concentrates in fewer technology categories.

These empirical evidences suggest that a firm innovation model has to explain a series of questions. Why do pathfinder and path-receiver patents have higher quality than average? Why are firms' future growth dependent on dynamics of the knowledge networks? How does the positive growth momentum spillover to peer firms? How do knowledge flows across sectors aggregate up?

We then build a tractable firm innovation model, featuring an endogenous network formation process at the patent level, that connects patent quality, citation formations and innovating firm performances. We describe the formation of a patent citation network as a node copying process with random copying errors, which is mathematically equivalent to a hybrid of the preferential attachment mechanism and a partially random formation mechanism, similar to Jackson and Rogers (2007), among others. Firm innovation results in new patents as new nodes (vertices) in the network. New patents cite existing ones and hence form new outgoing edges. Each patent has an *endogenous quality* that depreciates over time. A new patent exactly copy a fixed share of parent patent citations, erroneously copy another fixed share of parent patent citations with mutation to a random patent in the same cited sector, and randomly attach the rest of citations to any other existing patents with probability proportional to cited patent quality. A new patent of higher quality allocates a greater proportion of citations to random attachment and smaller share to follow parent's citations. An existing patent of high quality and many existing citations (indegree) is more likely to attract citations from new patents, either by being chosen as parent patent or through gaining random attached citation.

As a result, a new patent of high quality in one category is more likely to form edges with high-quality patents in a never-before-cited category through random attachment. That is why the expected patent quality of pathfinder (citing) and path-receiver (cited) patents are higher than average patent, which corresponds to Fact 1 of the empirical section. Once a pathfinder patent is granted by the patent office, the news reveal the quality information about both citing and cited patents, hence raise both pathfinder firm and path-receiver firm's profit and market value. This rationalizes pathfinder patent's impact on pathfinder firm and path-receiver firm in Facts 3 and 4 of the empirical section.

The mutation probability when copying parent patents explains the positive externality to peer firms that innovate in the same sector as the path-receiver patent, because future new patents will follow the new cross-sector citation path established this year, but the future new citations may deviate from the exact path-receiver patent and land in any other patents in the same cited sector. Peer firms in the pathfinder patent's sector (citing sector) also benefit because their new patents can choose the pathfinder patent as parent, follow the newly discovered citation path, and absorb knowledge from a new sector. These together explains the positive externality to peer firms in Facts 3 and 4 of the empirical section.

As new patents are born with no citations, it is the endogenously chosen patent quality that ultimately determines the citation dynamics in expectations. The law of motion of a patent's expected citation has a convenient closed-form solution. Moreover, the cross-sectional distribution of patent citations has an empirical consistent power-law right tail, as a standard preferential attachment model predicts, which we use to discipline model parameters in the calibration. The evolution of patent citation dynamics acts as demand function of knowledge in firm's innovation decision.

We group patents by sector and bundle the citation flows accordingly to produce an aggregated citation network. The aggregated network has a fixed number of nodes and growing counts of edges. Sectors with large knowledge stocks tend to attract as well as make more citations. Between two sectors, the form of the likelihood resembles gravity model, proportional to the product of patent numbers in both sectors. This aggregation result is consistent with our finding of Fact 5. Moreover, child-parent connections in the model generate a "home bias" that new patents tend to cite more existing ones in the same sector, consistent with the data.

Firm decisions in this model remain in line with the literature exemplified by Klette and Kortum (2004) and Lentz and Mortensen (2008). The main difference is that when making innovation decisions, forward looking firms take into consideration the future citation dynamics of their new patents. Firms need to decide on an additional margin — *patent quality* — that drives the citation dynamics and affects the demand for the corresponding product. High quality is costly but it generates profits on the product market and attracts future citations. When an existing patent gets a citation from a new patent, the owner firm of the existing patent receives a payment from the owner firm of the new one. The citation payment captures the application value of existing knowledge in a reduced form.

The new margin of firm innovation decision generates a *positive* impact of new knowledge on older patents, which we call the *innovation complementarity effect*. Existing knowledge has application value to new patents. Firms benefit from such application value when the older patents they own get new citations. A faster pace of aggregate knowledge growth increases the chance of an old patent receiving new citations, and it therefore increases the innovation incentives for forward looking firms as new knowledge now brings citation payments in the future. Owner firms of older patents still lose their market shares due to increased competition when new patents and hence new products are born, which is in line with the conventional creative destruction effect. Whether the net effect of aggregate knowledge growth on individual firm innovation decision is positive or negative depends on which effect dominates.

Firm's optimal innovation decision follows a cutoff rule, firm f innovates in sector j at t only when its innovation efficiency is sufficiently high. Firms with larger technology spaces are more likely to draw at least one innovation efficiency above the threshold, and then produce patents with high realized quality. Therefore, these firms are bigger and more likely to become pathfinders, which generates Fact 2 of the empirical section.

We assume that innovation is not purely random such that firms make sector-specific innovation decisions. This assumption reflects the observation that firms tend to specialize in their technology spaces. For example, a pharmaceutical firm is more likely to innovate in medicine related fields than in mining. A consequence of this assumption is that firm growths, even growths of firms with large knowledge stocks, can be volatile. When firms specialize in a handful of sectors, the within-firm Herfindahl indices are high, which capture the knowledge stock concentration within each firm. This agrees with Fact 6 of the empirical section. We call it the within-firm granularity, as a firm-level analogy to the economy-wide granularity by Gabaix (2011).

With the structured model that connects knowledge network formation and growth, we use Generalized Method of Moments (GMM) to estimate the model parameters by bringing data into general equilibrium conditions in the model. Some key model parameters change over time, such as average patent quality, number of citations per patent, within sector mutation rate of citation, royalty payment per citation, fixed cost of R&D per period, and marginal cost of increasing patent quality. We observe in the patent data that the quality adjusted growth rate of patents declines over time, which happens simultaneously with the productivity slow down of aggregate economy. Can our model help us understand the causes of innovation slow down? We employ counterfactual analysis by setting some model parameters at their 1976 initial level and solving the quality adjusted growth rate predicted by model, see if the counterfactual growth rate can restore 1976's high level. The culprit that counterfactual analysis finds out is the rising cost of R&D in the following three aspects: number of citations per patent, marginal cost of patent quality and research wage rate. Rising number of citations per patent may reflect the increasing rent protection effort in Dinopoulos and Syropoulos (2007) that increases that cost of new innovation. The rising cost of research quality agrees with the empirical evidence found in Bloom, Jones, Van Reenen, and Webb (2020) that research productivity has been declining in many sectors, such as Moore's Law, agriculture and medical innovations.

At the end of this paper, we summarize the model's policy implications according to firm level empirical evidences and results of counterfactual exercises. First, pathfinder patents build new connections in the knowledge networks, where firms absorb knowledge and receive reward of high quality patent. Not only the pathfinder firms and path-receiving firms benefit from the invention of such patents, peer firms which innovate in nearby technology space also enjoy positive spillovers. Therefore R&D subsidies to this type of high quality research need to be comparable to its positive externality. Second, R&D and IPR policies that reduce the rising cost of R&D can help reinstall high innovate rate, especially the number of citation per patent, which measures the cost of knowledge input in research. Note that the patent system has stopped adding new technology categories for decades. When the existing technology space get more crowded as more and more new inventions appear, each patent makes a smaller contribution and contains less intrinsic knowledge. It also makes new patents harder to prove originality. New patents are forced to cite more related prior arts to show respect and prevent future dispute. The larger quantity of knowledge stock in narrowly defined space may bring more burden than knowledge spillovers to new innovations. We share the same opinion as Boldrin and Levine (2013) that the IPR policy need to encourage the emergence of brand new technology categories that enlarge the dimension of technology space, instead of protecting large stakeholders in existing technology space. Third, we also find that average quality of patents, measured by the ability to attract random citations, has been declining since 1976. The model attributes this fact to the increasing marginal cost of patent quality in the innovation production function. Therefore, we suggest R&D subsidy target high quality research programs that tend to be more risky and time consuming than average projects. Lastly, immigration policy that welcomes high skilled workers provides abundant labor supply to R&D sector and keep research wage cost down.

## 2 Literature

This paper stands on the shoulders of four strands of literature. Firstly, more and more research focus on the dynamic formation process of the production and input network into the production network, such as, Bak, Chen, Scheinkman, and Woodford (1993), Vázquez (2003), Jackson and Rogers (2007), Chaney (2014), Oberfield (2018), Elliott, Golub, and Jackson (2014), Carvalho and Voigtländer (2014), Baqaee (2018), König (2016), Acemoglu and Azar (2017), Grassi et al. (2017), Taschereau-Dumouchel (2017) and Lim et al. (2017), etc. Among them, Vázquez (2003), Jackson and Rogers (2007), Chaney (2014) and Carvalho and Voigtländer (2014) assume that existing connections in the network help nodes establish new connections. This paper focuses on how citation network formation is related to quality of both citing and cited patents. A pathfinding patent is the first of its kind that connects knowledge in two previously isolated technology categories, therefore, our pathfinder patent measure of patent quality is close to the interdisciplinary and cutting edge measures in Higham, de Rassenfosse, and Jaffe (2020). This paper's theoretical model is therefore customized to explain the high quality of pathfinding patents and their positive externality to peer firms, adding quality based preferential attachment to standard dynamic network formation model. Many papers in the network formation literature are pure theoretical endeavour, this paper also provide empirical evidence on knowledge network dynamics and firm performance indices, such a market value, productivity, profit and innovation rate.

Secondly, this paper speaks to the literature of cross-sector knowledge diffusion. For example, Griliches (1957), Helpman (1998) and Jovanovic and Rousseau (2005) focus on General Purpose Technology (GPT), such as the electricity and semiconductor industry, the technologies of these industries have been widely used by other industries, creating an era of high growth rate and have become long-term growth engines. Akcigit, Celik, and Greenwood (2016) study the role of knowledge network in cross-firm transaction of intellectual property rights. Cai and Li (Forthcoming) study how cross-sector knowledge networks determine firms' expansion path across-sectors and the sectoral R&D intensity. Huang et al. (2018a) and Huang et al. (2018b) also pay attention to knowledge spillovers across industries and apply them to optimal R&D policy design and corporate mergers and acquisitions. Cai, Li, and Santacreu (2019) study the interaction between cross-country-sector knowledge spillovers and trade affect endogenous formation of comparative advantage and welfare gain from trade. They measure cross-sector knowledge spillovers by cross-country-sector patent citations. As far as we know, there is no empirical investigation to examine the link between knowledge diffusion networks and input and output table.

Third, this paper adds to the literature on firms market value and their patenting activities. See (Pakes and Schankerman (1984), Austin, 1993; Hall, Jaffe, and Trajtenberg (2001), Hall, Jaffe, and Trajtenberg (2005); Nicholas, 2008 andKogan, Papanikolaou, Seru, and Stoffman (2017)). These papers try to measure the economic value of patents. Some of them use citations received as a measure of patent quality. Our paper is closer to Kogan, Papanikolaou, Seru, and Stoffman (2017), which use the stock market reaction after patent announcement as a means to construct appropriate measure of patent quality to study within- as well as between-industry reallocation and growth dynamics. They find positive knowledge spillovers or business stealing effects to peers as well as in Bloom, Schankerman, and Van Reenen (2013). Our pathfinder patent predicts higher firm performance measures for both the citing and cited firm involved; it also brings positive externality to peer firms in nearby technology space of the new citation link. Our work are also related to the papers that study the impact of innovation on firm productivity and growth (Caballero and Jaffe, 1993; Akcigit and Kerr, 2010; Acemoglu, Akcigit, Alp, Bloom, and Kerr (2018)).

Fourth, this paper is related to the literature that construct and compare different patent quality measures summarized by Higham, de Rassenfosse, and Jaffe (2020), starting from Trajtenberg, Henderson, and Jaffe (1997) to Higham, Governale, Jaffe, and Zülicke (2019), and Marx and Fuegi (2020). Since we define high quality patents as being more likely to cite and be cited by other high quality patents through random attachment, instead of following parent citations, our quality definition is close to forward citation count but excluding citations from relative patents. Our definition of pathfinder patent is similar to the identification of the first patent that used a novel word in the patent application text in Balsmeier, Assaf, Chesebro, Fierro, Johnson, Johnson, Li, Lück, O'Reagan, Yeh et al. (2018) in the sense of finding the very first patent of its kind. The definition of pathfinder patents by nature is also related to interdisciplinary and cutting edge measures in Higham, de Rassenfosse, and Jaffe (2020), because they build new links between previous disconnected patent classes.

The rest of this paper proceeds as the following. In Section 3, we describe the Patent data we use, compare pathfinder patents and pathfinder firms with peers, and present firm level evidences on pathfinder patent's impact on pathfinder firms, path-receiver firms and peer firms. In Section 4, we build the firm innovation model with dynamics of knowledge networks that explain the empirical facts found in previous section. In Section 5, we calibrate the model and run counterfactual analysis to detect the factors that contribute to quality weighted innovation rate slow down from 1976. Lastly, we provide policy implications and conclude.

# 3 Empirics

In this section, we first describe the patent database and matched patent data with Compustat data. Then we present several empirical facts related to pathfinder, path-receiver firms and peer firms' real performances in response to knowledge network dynamics, citation flow across technology class and firm growth volatility.

## 3.1 Data

**USPTO Patent Data.** The empirical analysis draws from the USPTO Patent Database 2015 version. The dataset contains detailed information on 5.9 million utility patents granted by the U.S. Patent and Trademark Office between the years 1976 and 2015. A patent has to cite another patent when the former has content related to the latter. When patent A cites patent B, this particular citation becomes both a backward citation made by A to B and a forward citation received by B from A. Moreover, the patent data contains US patent class code for each patent that helps identify where it lies in the technology space.

We use patent citation data to identify if a patent A in category i that cites a patent B in category j at year t is the first patent that cites from i to j in the history. We call A a pathfinder patent and B a path-receiver patent. Then we find A (B)'s owner firm and call it a pathfinder (path-receiver) firm in year t. Other firms that do not own any pathfinder patent nor path-receiver patent in year t are called peer firms.

**Compustat North American Fundamentals (Annual).** In order to assess the impact of patents and their technological distance on firm moments, such as stock market valuation, the PDP patent data is linked to Compustat firms. The focus is on the balance sheets of Compustat firms between the years 1974-2006, retrieved from Wharton Research Data Services. The Compustat database and the NBER PDP database are connected using the matching procedure provided in the PDP data. When testing if market value changes with patent grants, we use matched US Patent and

Compustat North American Fundamentals dataset provided by Kogan, Papanikolaou, Seru, and Stoffman (2017). Firm level total factor productivity is taken from İmrohoroğlu and Tüzel (2014).

### 3.2 Pathfinder Patents and Firms

We find that pathfinder patents score higher in several quality measures in the literature, such as originality, generality (Trajtenberg, Henderson, and Jaffe (1997)), number of forward citations, number of Collaborated Patent Classification (CPC) technology class memberships, backward-citations' pedigree, average age of backward citations, grant lag, and number of backward citations to scientific literature used in Higham, de Rassenfosse, and Jaffe (2020).[1] Citation to scientific literature data is taken from Marx and Fuegi (2020).

More over, pathfinder and path-receiver firms are much larger than average firms in the patent data base, both in terms of total patent counts and number of patent categories. This fact helps explain in the coming facts, why citation network dynamics some times have stronger impact on peer firms, that are much smaller, than on super large pathfinder and path-receiver firms.

**Fact 1.** Path-finder patents receive more forward citations, have higher generality and originality score, wider CPC technology class memberships, lower backward-citations' pedigree, cite younger patents, shorter grant lags, more backward citations to scientific literature and fewer novel word than average patents.

| | All Patents | Pathfinder Patents | Pathreceiver Patents |
|---|---|---|---|
| Originality | 0.52 | 0.79 | 0.59 |
| Generality | 0.55 | 0.64 | 0.72 |
| # Forward citations received | 11.05 | 17.24 | 22.49 |
| # CPC technology class memberships -1 | 0.94 | 1.10 | 1.01 |
| Backward-citations' pedigree | 1.60 | 1.38 | 1.48 |
| Average age of backward citations | 11.81 | 9.90 | 9.12 |
| Grant lag | 2.63 | 2.51 | 2.15 |
| # Backward citations to scientific literature (front page) | 14.86 | 18.82 | 3.96 |
| #Novel word per patent | 0.41 | 0.29 | 0.21 |

In most quality measures, pathfinder patents have higher quality than average as defined by the literature, except for backward-citations' pedigree, grant lag and novel word count. For example, Higham, de Rassenfosse, and Jaffe (2020) think lower backward-citations' pedigree and shorter grant lag predict lower quality of patents. Balsmeier, Assaf, Chesebro, Fierro, Johnson, Johnson, Li, Lück, O'Reagan, Yeh et al. (2018) use weighted novel word count to evaluate the originality of a patent. The discrepancy happens because pathfinder patents' novelty comes from neither citing other popular patents (as captured by backward-citations' pedigree) nor being the first to create a new idea (novel word); instead, pathfinder patents are novel because they use a new combination of knowledge or material inputs to produce an existing product in a better way.

[1] Backward-citations' pedigree measure is granted from Higham, Governale, Jaffe, and Zülicke (2019). We thank Kyle Higham for sharing the pedigree data with us.

Moreover, the novelty of pathfinder patents is also obvious to judge, hence they are quicker to be granted.

Another interesting finding is that path-receiver patents have similar quality measures as pathfinder patents, except for number of backward citations to scientific literature. Since backward citations to scientific literature measures a patent's relation to frontier scientific research, lower score in this measure means path-receiver patents are relatively well established and easy to be understood by outsiders compare to average patents. This is reasonable, because when inventors utilize knowledge from a new sector for the first time, they start from those simple and tested patents, instead of fancy and frontier ones.

**Fact 2.** Pathfinder and path-receiver firms own greater number of patents and innovate in larger set of technology categories than average firms.

| | All Patenting Firms | Pathfinder Firms | Path-receiver Firms |
|---|---|---|---|
| Mean Patent stock | 45.72 | 250.86 | 451.30 |
| Mean Patent categories | 1.84 | 5.99 | 7.75 |

### 3.3 Citation Network Dynamics and Firm Innovation

**Fact 3.** A firm's innovate rate is higher in future 5 year, when there are more new cross-category inward and outward links made on firm's technology space either by self firm or peer firms, especially so for firm's extensive growth into new technology categories.

In this subsection, we test in the patent data if firms' rate of innovation changes, when there are new cross-sector citation links made to or from the patents categories that they have patented before. First, we find that firms innovate faster, when they made new cross-category citation links by themselves. Then, we further discover that firms' innovation rate receives externalities from new citation links made by peer firms.

$$\begin{aligned} log(\frac{ps_{f,t+\tau}}{ps_{f,t}}) =& \beta_{p,\tau} \log(ps_{f,t}) + \beta_{n,\tau} \log(nonclass_{f,t}) + \\ & \beta^{self}_{pf,\tau} \log(wpf^{self}_{f,t} + 1) + \beta^{peer}_{pf,\tau} \log(wpf^{peer}_{f,t} + 1) + \\ & \beta^{self}_{pr,\tau} \log(wpr^{self}_{f,t} + 1) + \beta^{peer}_{pr,\tau} \log(wpr^{peer}_{f,t} + 1) + \\ & D_{f,\tau} + D_{t,\tau} + D_{ind,\tau} + D_{t,\tau} + \epsilon_{f,t,\tau} \end{aligned} \tag{1}$$

where $\tau$ is the growth time horizon in years, $ps_{f,t}$ is the total number of patents granted to firm f by time t; $nonclass_{f,t}$ is the number of technology categories that firm f has patent application at time t, it measures firm f's patent scope; $wpf^{self}_{f,t}$ is the patent stock weighted number of new outward cross-category citations links made by firm f at time t (or paths found), it measures firm f's ability to find new knowledge source and make high quality innovation; $wpr^{self}_{f,t}$ is the patent stock weighted number of new inward cross-category citation links received by firm f at time t (or paths received), it measures firm f's ability to attract citation from and find knowledge application at a new source; $wpr^{peer}_{f,t}$ is the weighted number of inward citation links made to firm f's technology space made by peer firms, it measures firm f's technology space's ability to find applications in new areas; and $wpf^{peer}_{f,t}$ is the weighted number of outward citation links made to firm f's technology space made by peer firms, it measures firm f's technology space's strength in obtaining knowledge from new sources. $D_{f,\tau}$, $D_{ind,\tau}$ and $D_{t,\tau}$ are firm, industry and year time dummies. The following equations give calculation details of these variables.

$$wpr^{self}_{f,t} = \sum_{j=1}^{J} \frac{ps^{j}_{f,t}}{ps_{f,t}} newindegree^{self}_{j,t}, \tag{2}$$

$$wpf^{self}_{f,t} = \sum_{j=1}^{J} \frac{ps^{j}_{f,t}}{ps_{f,t}} newoutdegree^{self}_{j,t}, \tag{3}$$

$$wpr^{peer}_{f,t} = \sum_{j=1}^{J} \frac{ps^{j}_{f,t}}{ps_{f,t}} newindegree^{peer}_{j,t}, \tag{4}$$

and

$$wpf^{peer}_{f,t} = \sum_{j=1}^{J} \frac{ps^{j}_{f,t}}{ps_{f,t}} newoutdegree^{peer}_{j,t} \tag{5}$$

where $newindegree^{self}_{j,t}$ is the number of new inward cross-category citation links made to category j at time t by firm f itself, $newoutdegree^{self}_{j,t}$ is the number of new outward cross-category citation links made from category j at time t by firm f itself. $newindegree^{peer}_{j,t}$ is the number of new inward cross-category citation links made to category j at time t by other firms, $newoutdegree^{peer}_{j,t}$ is the number of new outward cross-category citation links made from category j at time t by other firms.

$ps^{j}_{f,t}$ is firm f's number of patent stock in patent category j by time t. $ps_{f,t}$ is firm f's total number of patent stock in all categories by time t. Even for firms that compete in the same product market, they may still own various combinations of patent stocks as documented by Bloom, Schankerman, and Van Reenen (2013). $\{\frac{ps^{j}_{f,t}}{ps_{f,t}}\}$ represents firms' heterogeneity in their portfolio of patent stocks across different technology categories, we use it to measure firms' various exposure to citation network dynamic changes at a given time t.

We present the regression results in Tables 2, 3 and 4. We find that the ability to make or receive new cross-category citation links gives firms a positive long lasting boost in future innovation. Moreover, the number of new citation links received $lwpr_{self}$ has much stronger impact on growth rate than number of citation links sent out $lwpf_{self}$. Maybe because $lwpr_{self}$ is a stronger indicator of firm f's patents' quality and potential applications than $lwpf_{self}$. When peer firms made or received new links, firms also also innovate faster in the future, even though the impact is one order smaller than that of self made or received new links.

To find out the mechanism through which knowledge networks dynamics affects firm innovation rate, we split firm growth rate in patent number into extensive growth and intensive growth. Extensive growth captures a firm's patent application growth in new technology categories that this firm has never patented before. Intensive growth measures a firm's patent application growth in existing categories. In Table 3, we show that both inward and outward new links that happened to own or peer firms predict long lasting higher future extensive growth rate at firm level. However, in Table 4, firms's intensive growth reacts negatively to peer's new links of both types, even to own firm's inward link after 1 year. Both types of new links associated with own firms increase firm intensive growth rate for 1 or 2 years in the short run only. Overall, we conclude that new cross-category citation links promote firm innovation rate mainly by helping firms to expand into new categories.

### 3.4 Citation Network Dynamics and Firm Real Performances

**Fact 4.** A public listed firm's market value, employment, capital, sales, profit and TFP is larger in the future 5 years, when there are more new cross-category inward and outward links made on firm's technology space either by self firm or peer firms.

In this subsection, we test in linked patent and CRSP/Compustat data, if firms' stock market value, employment, capital, sales, profit and TFP changes, in response to new cross-category citation links made in the citation networks. First, we show that firms' market value, labor and capital inputs, profit, sales and TFP increase, when there are new cross-category citation links in their technology space made by themselves. Then, we find that firms receive consistently positive externalities from new cross-category citation links made by other firms.

$$
\begin{aligned}
log(\frac{X_{f,t+\tau}}{X_{f,t}}) =& \beta_{p,\tau} \log(ps_{f,t}) + \beta_{n,\tau} \log(nonclass_{f,t}) + \\
& \beta^{self}_{pf,\tau} \log(wpf^{self}_{f,t} * 10^5 + 1) + \beta^{peer}_{pf,\tau} \log(wpf^{peer}_{f,t} * 10^5 + 1) + \\
& \beta^{self}_{pr,\tau} \log(wpr^{self}_{f,t} * 10^5 + 1) + \beta^{peer}_{pr,\tau} \log(wpr^{peer}_{f,t} * 10^5 + 1) + \\
& \log(tsm_{f,t}) + \log(tcw_{f,t}) + Z_{f,t} + D_{f,\tau} + D_{t,\tau} + D_{ind,\tau} + \epsilon_{f,t,\tau},
\end{aligned} \tag{6}
$$

where $X_{f,t}$ can be market value, total employment, capital stock, profit, sales and TFP[2] of firm f at time t. $tsm_{f,t}$ is total dollar value of innovation produced by firm f in year t, based on stock market (sm); and $tcw_{f,t}$ is a measure of the output of innovation produced by a firm using its citation-weighted (cw) patents. We take these two measures from Kogan, Papanikolaou, Seru, and Stoffman (2017). $Z_{f,t}$ is the set of size controls, such as log scaled firm f's total employment, capital and $X_{f,t}$ itself at time t (if $X_{f,t}$ is not employment or capital). $D_{f,\tau}$, $D_{ind,\tau}$ and $D_{t,\tau}$ are firm, industry and year time dummies. All variables are standardized so that coefficients are

[2] Taken from İmrohoroğlu and Tüzel (2014)'s web posted firm TFP data. See https://sites.google.com/usc.edu/selale-tuzel/home?authuser=2

comparable across them. The coefficients now means when the correspondent variable increases one standard deviation, how many percent growth rate in $X_{f,t}$ changes.

We present our regression results in Tables 5 to 10. In Table 5, a firm's market value in stock market increases 1.88-4.64% in the first year, when this firm is granted one standard deviation more pathfinder patents; market value increases 0.33-0.40% in the first year, when other firms are granted one standard deviation more pathfinder patents. Likewise, a firm's market value in stock market increases 6.1% in the first year, when this firm's patents received one standard deviation more pathfinder citations from pathfinder patents; but we do not find significant externality when peer firms receive pathfinding citations. When we control industry fixed effects only, the knowledge network dynamics' impact on firm value can persist after up to 5 years, with gradually fading coefficients. However, when we control firm fixed effect, the positive impacts only exist for the first year and only for pathfinder patents.

In Tables 6 and 7, firms' production inputs, both labor and capital, increases in the range of 1-4% in response to one standard deviation increase in the grant of pathfinder patents to own firm and peer firms. The positive impacts are also long lasting, in many cases coefficients increase over time. The impact of own pathfinder patents tend to be larger than peer firm pathfinder patents. In Tables 8 and 9, we observe that firms' profit and sales behave in similar manner as labor and capital inputs. In Table 10, we find that firms' TFP response positively to own pathfinder and path-receiver patents, but not to peer pathfinder and path-receiver patents.

Overall, we find that dynamics in the knowledge networks have positive real impact on firm market value, labor and capital inputs, sales, profit and TFP, not only to the pathfinder firms themselves, but also to peer firms who innovate in the neighbourhood technology space. In the Appendix, we conducted robustness checks using alternative measures of firms' exposure to knowledge network dynamics. New citation links every year are weighted by edge betweenness centrality of a new link to capture the various importance of new citation links to the network topology. Similar empirical results still hold in Tables 13 and 18.

### 3.5 Citation Flows across Categories

**Fact 5.** Citation counts across technology categories follow a gravity model, which is positively related to the patent stocks in both categories.

In the following two equations, we regress citation flow between citing and cited patent categories on the patent stocks of citing category and cited category, controlling different combinations of fixed effects.

$$\log(citation_{cd,t}) = \beta_c \log(ps_{c,t}) + \beta_d \log(ps_{d,t}) + D_c + D_d + D_t \tag{7}$$

$$\log(citation_{cd,t}) = \beta_c \log(ps_{c,t}) + \beta_d \log(ps_{d,t}) + D_t + D_{cd} \tag{8}$$

In Table 11 we show that citation flow between sectors follows a gravity type model that is widely used to study trade and immigration flow across countries and sectors.

### 3.6 Firm Growth Volatility and Innovation Concentration across Sectors

**Fact 6.** Firm growth volatility is larger when the firm's innovation concentrates in a few technology categories.

We measure growth volatility $g\_vol5$ using variance of growth rates in the future 5 years and innovation concentration with the Herfindahl index of patent share across categories.

$$Herfindahl_{f,t} = \sum_{j=1}^{J} (ps^{j}_{f,t}/ps_{f,t})^2 \tag{9}$$

Then adjust for the number of categories $nonclass_{f,t}$,

$$Herfindahl^{a}_{f,t} = \frac{Herfindahl_{f,t} - 1/nonclass_{f,t}}{1 - 1/nonclass_{f,t}} \tag{10}$$

Then we run the following regression and present the result in Table 19. Firms with larger patent stock experience lower growth volatility, while firms with more concentrated patent portfolio have more turbulent growth paths.

$$\log(g\_vol5_{f,t}) = \beta^{v}_{p} log(ps_{f,t}) + \beta^{v}_{n} log(nonclass_{f,t}) + \beta^{v}_{h} Herfindahl^{a}_{f,t} + D_f + D_t \tag{11}$$

## 4 Theoretical framework

We rationalize our empirical findings in a firm innovation model featuring the endogenous dynamics of a patent citation network.

Time is continuous with an infinite horizon. The economy is populated by a unit-mass continuum of identical price-taking households who work and consume. The economy has a fixed continuum $\mathcal{F}$ of firms with mass $F > 0$. Firms hire in a perfectly competitive labor market, and they produce and innovate. Let $\mathcal{J}$ be the set of all $J$ sectors.[3] Define the *technology space* of a firm $f$ as $\mathcal{J}^f \subseteq \mathcal{J}, \forall f \in \mathcal{F}$. Firm innovation entails developing new patents over time in each sector of the firm's technology space.[4] Assume that every sector lies in the technology spaces of a continuum of firms. A patent corresponds to a product variety and its production technology. Only the owner firm of a patent can produce and sell the corresponding product. Product markets are monopolistically competitive.

[3] Each sector $j \in \mathcal{J}$ can be viewed as one technological category or a set of them in the data.

[4] We fix $\mathcal{J}^f$ for each firm $f \in \mathcal{F}$, as we abstract away from firms' decisions to expand their technology space or to exit the economy.

The set of available product varieties for household consumption is therefore endogenous, denoted as $\mathcal{N}(t)$ at any time $t$, which also represents all existing patents at $t$. Every patent has an industry classification $j \in \mathcal{J}$ and belongs to a firm $f \in \mathcal{F}$. At time $t$, $n_j^f(t)$ is the number of firm $f$'s patents in industry $j$, and the number of sector-$j$ varieties $f$ produces, such that $n_j^f(t) = 0$ if $j \notin \mathcal{J}^f$. Firm $f$'s *knowledge stock* is given as $n^f(t) = \sum_{j \in \mathcal{J}^f} n_j^f(t)$; sector $j$ has $N_j(t)$ available varieties at $t$, with $N_j(t) = \int_{f \in \mathcal{F}} n_j^f(t)$. The aggregate knowledge stock $N(t)$ is the cardinality of $\mathcal{N}(t)$, satisfying $N(t) = \sum_{j \in \mathcal{J}} N_j(t) = \int_{f \in \mathcal{F}} n^f(t)$.

In what follows, subsection 4.1 describes the simple problem of the households. Then, subsection 4.2 introduces a network formation model of patent citations, where the network is a growing digraph at the *patent level*, with each vertex being a patent and each directed edge a citation. Next, subsection 4.3 and subsection 4.4 discuss firm decisions in detail, where firms rationally anticipate the patent-level network dynamics. We close the model in subsection 4.5 and define an equilibrium. The analytical form of a firm's value function is in subsection 4.6. Firm and sectoral growths are discussed in subsection 4.7. Lastly, subsection 4.8 considers network aggregation.

### 4.1 Household preferences

The identical households discount the future at rate $\rho > 0$. A representative household can fully summarize their behavior, with the following life-time preferences:

$$U = \int_0^\infty e^{-\rho t} \log C(t) \mathrm{d}t,$$

where $C(t)$ is the consumption compound such that it aggregates all available varieties with constant elasticity of substitution (CES), given as

$$C(t) = \left( \int_{\omega \in \mathcal{N}(t)} z(\omega, t)^{\frac{1}{\nu}} c(\omega, t)^{\frac{\nu - 1}{\nu}} \mathrm{d}\omega \right)^{\frac{\nu}{\nu - 1}},$$

where $c(\omega, t)$ is the consumption of variety $\omega$ at $t$, $z(\omega, t)$ is the *quality* of variety $\omega$ at $t$, and $\nu > 1$ is the substitution elasticity among varieties. Quality $z(\omega, t)$ results from each firm's endogenous choice, which the households take as given. Household income consists of labor income and profit payments as firm owners. Labor is in fixed supply and consists of production labor $\overline{L} > 0$ and research labor $\overline{R} > 0$. Henceforth, we choose the production labor as the numeraire and denote the wage rate of research labor as $w_R(t)$. We look for an equilibrium with a fixed research wage rate $w_R(t) = w_R$ and a fixed aggregate profit generated by all firms $\Pi(t) = \Pi$, $\forall t$. The representative household's budget constraint at any $t$ then reduces to $\int_{\omega \in \mathcal{N}(t)} p(\omega, t) c(\omega, t) \mathrm{d}\omega = \overline{L} + w_R \overline{R} + \Pi$.

Given the qualities and market prices of all varieties $\{z(\omega, t), p(\omega, t)\}_{\omega \in \mathcal{N}(t)}$, standard cost minimization yields the demand function of each product variety such that

$$c(\omega, t) = z(\omega, t) C(t) \left( \frac{p(\omega, t)}{P(t)} \right)^{-\nu}, \tag{12}$$

where $P(t)$ is the ideal price index, given as

$$P(t) = \left( \int_{\omega \in \mathcal{N}(t)} z(\omega, t) p(\omega, t)^{-(\nu-1)} \mathrm{d}\omega \right)^{-\frac{1}{\nu-1}}.$$

The market clearing condition for each $\omega$ at $t$ is $y(\omega, t) = c(\omega, t)$, where $y(\omega, t)$ is the output level.

### 4.2 Dynamics of the patent citation network

We describe the dynamics of the patent network as the continuous limit of a growing network with discrete vertices, in the same spirit of the "mean-field" approximation by Jackson and Rogers (2007). At any time, the set of all patents and the citation flows among them form a digraph, where each patent is a vertex or node and each citation is a *directed* edge from the citing patent to the cited patent.[5] The network grows over time as firm innovation results in new nodes as well as new edges. An existing patent gets new cites only at the birth of new patents; symmetrically, only the newborn patents add new citation edges to other patents. All citation edges are permanent and cannot be removed or rewired once established.

The goal of this section is to describe the evolution dynamics of patent citations. More specifically, we look for two things. One is the law of motion of the number of citations a patent is expected to receive over time. As will become clear in subsection 4.4, this law of motion is payoff-relevant to innovating firms. It will serve a similar role to the product demand in eq. (12) as a "demand curve" for a firm's new knowledge in the firm's decision. The other is the cross-sectional distribution of citations. Following standard practice, we focus on a stationary cross-sectional citation distribution in this growing network. This distribution will confront its data counterpart, and help to discipline model parameters.

Suppose that the network grows at a constant rate $g > 0$, such that the increment in network size from $t$ to $t + \mathrm{d}t$ satisfies $\frac{N(t+\mathrm{d}t)}{N(t)} = 1 + g\mathrm{d}t + o(\mathrm{d}t)$ for small $\mathrm{d}t$. The initial size $N(0) = N_0$ is given. For any node $\omega \in \mathcal{N}(t)$, define its *indegree* as the total number of inward edges (citations received) by $t$ and its *outdegree* as the number of outgoing edges (citations made). A node's indegree may increase over time whereas its outdegree is fixed. Let $\bar{x} > 0$ be the exogenous and constant number of citations a new patent is expected to make, then $\bar{x}$ is the average in- or out-degree of the network at any time.

Each patent has an intrinsic feature $z$ upon birth, referred to as the *quality*, such that $z \in [z_{\min}, z_{\max}]$ with $0 < z_{\min} < z_{\max} < 1$. Over time, quality depreciates at an exogenous rate $\delta > 0$. If a patent $\omega$ has an initial quality $z$ upon birth at time $t$, then its quality at age $\tau$ is $z(\omega, t + \tau) = ze^{-\delta\tau}$. Quality depreciation reflects knowledge obsolescence over time. It also implies that, *ceteris paribus*, households prefer newer varieties. Let each new patent's initial quality $z$ be random draws from a fixed distribution $Z$ over support $[z_{\min}, z_{\max}]$ with mean $\bar{z}$.

The network growth rate $g$ and the quality depreciation rate $\delta$ jointly determined the average

[5] In this paper, we use "vertex" and "node" interchangeably, as well as "edge" and "link".

quality $\overline{z^o}(t)$ of all existing ($o$ for old) patents. Without loss of generality, suppose that at time 0, the average quality is $\overline{z^o}(0) = \bar{z}\frac{g}{g+\delta}$. Then, at any time $t$, the average quality $\overline{z^o}(t)$ remains unchanged at $\overline{z^o} \equiv \bar{z}\frac{g}{g+\delta}$. This is because $\overline{z^o}$ is the limit of average quality as $t \to \infty$, such that

$$\overline{z^o} = \int_0^\infty g e^{-g\tau} \bar{z} e^{-\delta\tau} \mathrm{d}\tau = \bar{z}\frac{g}{g+\delta},$$

where $ge^{-g\tau}$ is the limit cross-sectional probability density function (PDF) of patent age $\tau \geq 0$. All sectors share the common $\overline{z^o}$.

At this stage, it suffices to treat the initial quality distribution $Z$, its mean $\bar{z}$, and the network size growth rate $g$ as exogenous. We endogenize them in subsection 4.4 as the equilibrium outcome.

Matching between every new patent and existing ones to form new citation links follows an *imperfect node copying process*, where a new patent can replicate an existing one's citations with copying errors (*mutations*).[6] Figure 1 illustrates the four ways of forming new edges. When a firm innovates in a sector $j$, it creates new sector-$j$ patents based on existing ones in $j$. An existing node becomes the base upon which a new patent is developed at random, with probability proportional to the existing node's current quality. Then the existing node is referred to as the *parent* of the new one (*child*). A child must cite its parent. It is also expected to cite $\bar{x} - 1$ other existing nodes. Suppose the child has an initial quality $z$. The child cites $z\bar{x} - 1$ nodes that are neither its parent nor cited by its parent. The probability that each of these nodes is picked is proportional to their quality. The child distributes the remaining $(1 - z)\bar{x}$ citations by randomly replicating the parent's citations with equal probability. We allow the replication to *mutate* within a sector, with probability $\bar{\eta} \in [0, 1]$. Without mutation, the child cites the same node as the parent does; with mutation, the child cites an alternative node that is in the same sector of the node its parent cites. The alternative node is chosen with probability proportional to its quality.

Patent quality $z$ plays two more roles in the growth of a citation network, in addition to affecting the demand for the corresponding variety on the product markets. First, upon birth, a patent's quality determines its *originality* in the sense that a higher-quality new patent shares fewer similarities with its parent patent, reflected by a smaller set of common citations. Second, after birth, a patent's quality determines the expected rate at which it attracts citations from newer patents.

We are ready to characterize the expected growth of an arbitrary patent's citations, i.e., the indegree of a node. Consider an existing patent named $\omega$ in sector $j$ with initial quality $z$, age $\tau$, and current indegree $d$ at time $t$. This node's current quality is $z(\omega, t) = ze^{-\delta\tau}$. A new node is expected to have initial quality $\bar{z}$ and make $\bar{x}$ citations. Symmetrically to Figure 1, node $\omega$ receives new cites from a newborn node at $t$ in four ways. The first way is to become the parent node, with probability density $ze^{-\delta\tau}/[N(t)\overline{z^o}]$. This is because the new node belongs to sector $j$ with probability $N_j(t)/N(t)$, and then it finds node $\omega$ as its parent with probability $ze^{-\delta\tau}/[N_j(t)\overline{z^o}]$. The second way is via exact citation replication. Node $\omega$ connects to $d$ other existing ones, each of which has

[6] For a simpler textbook description of the node copying process, see section 13.5 of Newman (2018).

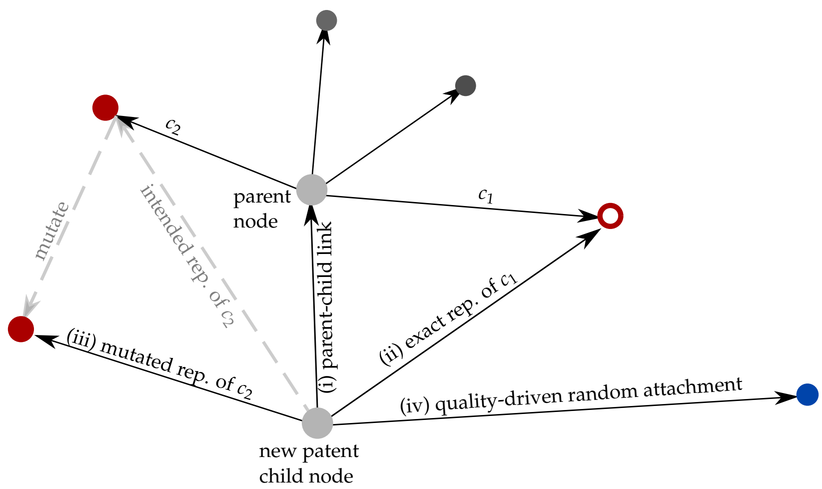


Figure 1: Four ways for a newborn node to form new edges: an illustration.
Each marker style represents patents in the same sector. A new node adds new edges to the network by (i) citing its parent node, (ii) replicating its parent's citations with (iii) possible mutations, and (iv) independently citing other existing nodes.

a $1/N(t)$ probability density of being the new node's parent, and the new node replicates this citation with probability $(1-\overline{z})(1-\overline{\eta})$. The third way is due to mutated replication. The new node is expected to make $\overline{x}(1-\overline{z})\overline{\eta}$ citations (outgoing edges) by mutated replications. Without mutation, each of such edges would go to a node in sector $j$ other than $\omega$ with probability $N_j(t)/N(t)$; and after mutation, the edge goes to $\omega$ with probability $ze^{-\delta\tau}/[N_j(t)\overline{z^o}]$. Combined, the expected number of new cites that node $\omega$ gets due to mutated replication is $ze^{-\delta\tau}\overline{x}(1-\overline{z})\overline{\eta}/[N(t)\overline{z^o}]$. The last way is that node $\omega$ gets original new cites untied to the new node's parent. Each of such cites occurs with probability density $ze^{-\delta\tau}/[N(t)\overline{z^o}]$, and the new node makes $\overline{xz}-1$ such citations in expectation. Therefore, when a new node is born at time $t$ following the imperfect node copying process, the expected number of new citations that node $\omega$ in sector $j$ with age $\tau$ and initial quality $z$ gets is as follows

$$\begin{aligned}&\frac{ze^{-\delta\tau}}{N(t)\overline{z^o}}+d\frac{(1-\overline{\eta})(1-\overline{z})}{N(t)}+ze^{-\delta\tau}\frac{\overline{x}(1-\overline{z})\overline{\eta}}{N(t)\overline{z^o}}+ze^{-\delta\tau}\frac{\overline{xz}-1}{N(t)\overline{z^o}}\\ =&\frac{1}{N(t)}\left[\frac{ze^{-\delta\tau}}{\overline{z}g/(g+\delta)}\overline{x}(1-\overline{r})+d\overline{r}\right],\end{aligned} \tag{13}$$

where, for convenience, $\overline{r}\equiv(1-\overline{\eta})(1-\overline{z})$ summarizes the probability of an *exact replication*. The intuition is that replication rate $\overline{r}$ determines the citation growth driven by existing citations $d$, and the complement rate $1-\overline{r}$ sets the rate at which a patent's relative quality attracts new citations. It is straightforward to verify that when a new node is born at time $t$, the total number of new citations that all existing patents $\omega\in\mathcal{N}(t)$ expect to receive is indeed $\overline{x}$. Notably, the network formation model we present here is equivalent to a hybrid mechanism of preferential attachment and quality-dependent random matching.

The law of motion of a patent's expected number of citations has an analytical solution. Let $k(z,\tau)$ be the unconditional expectation of indegree or number of citations a patent with initial quality $z$ has at age $\tau$. In particular, we have $k(z,0) = 0, \forall z$, because a newborn patent has not yet received any citations. Moreover, constant network growth means that $k(z,\tau)$ is independent of the node's time of birth.

**Proposition 1.** *(Law of motion of expected indegrees.) Consider a growing patent citation network where new patent nodes are born at an expected rate $g > 0$, following the imperfect node copying process described above. Regardless of the time of birth, a patent with initial quality $z$ expects to get $k(z,\tau)$ citations at age $\tau$, such that $k(z,\tau)$ solves a differential equation given as*

$$\partial_\tau k(z,\tau) = \frac{ze^{-\delta\tau}}{\overline{z}/(g+\delta)}\overline{x}(1-\overline{r}) + k(z,\tau)g\overline{r}, \tag{14}$$

*with initial condition $k(z,0) = 0, \forall z$. Notation $\partial_\tau k(z,\tau)$ is shorthand for $\frac{\partial k(z,\tau)}{\partial\tau}$. Therefore, the unconditional expectation of a node's indegree $k(z,\tau)$ at age $\tau$ is a linear function of its initial quality $z$, satisfying*

$$k(z,\tau) = ze^{-\delta\tau}\frac{\overline{x}(1-\overline{r})(g+\delta)}{\overline{z}(g\overline{r}+\delta)}\left[e^{(g\overline{r}+\delta)\tau} - 1\right]. \tag{15}$$

*It follows that $\partial_\tau k(z,\tau)$ is also linear and strictly increasing in $z$.*

In this proposition, eq. (14) shows that new citations are more likely attracted to existing patents that have relatively high current quality ($ze^{-\delta\tau}$) and are already widely cited ($k(z,\tau)$). Ultimately, it is the intrinsic quality of a patent that determines the level and growth of its expected number of citations, as eq. (15) shows. A patent with $z > 0$ accumulates citations as it ages; but the citation growth rate $\frac{\partial_\tau k(z,\tau)}{k(z,\tau)} = \partial_\tau(\log k(z,\tau))$ is strictly decreasing in age $\tau$, more so if quality depreciates at a higher rate $\delta$.

The network formation process helps the full model to capture Fact 1 and Fact 4. Suppose a patent $\omega_j$ in sector $j$ receives a citation from a new patent $\omega_i$ in sector $i$, and this is the first citation from $i$ to $j$. The new node $\omega_i$ is hence the pathfinder patent. Conditional on establishing a new path, $\omega_j$ is likely of high quality and so is $\omega_i$. The reason is that such an event rules out all the other ways for $\omega_i$ to cite $\omega_j$ except for random attachment driven by $\omega_j$'s high quality — two patents in different sectors cannot form parent-child link and neither exact nor mutated replication is possible for the first cross-sector citation. For the same reason, random attachment is more likely to occur when the citing patent $\omega_i$ is of high quality and hence makes more independent citations untied to its parent. This positive relation between being a pathfinder and having high quality is consistent with Fact 1. Consequently, if quality is payoff relevant, a pathfinding event is good news for both the pathfinder firm and the receiver firm, and their market value should increase. Meanwhile, the impact of a pathfinding event spills over to other firms due to mutation. Consider a firm $f_1$ that also innovates in $j$ but does not own $\omega_j$. When $\omega_j$ is born and cites $\omega_i$, the immediate impact on firm $f_1$ is negative due to diluted market share and a lower chance of getting cited in the future. However, there is a positive indirect effect. Now that the pathway from $j$ to $i$ is established, if a future patent of $f_1$ identifies $\omega_j$ as the parent, then it is can utilize the pathway through exact or mutated replication. The same reason applies to a firm $f_2$ that innovates in $i$ but does not own $\omega_i$. Eventually, the new pathway increases the chance of $f_2$'s patents in $i$ getting cited through mutated replication. These are consistent with Fact 4.

The next task is to characterize the stationary cross-sectional distribution of indegrees when the network size is sufficiently large. The focus is the shape of the distribution's right tail, when indegrees become larger. The benchmark is a Pareto distribution whose counter cumulative distribution function (CCDF, right tail) has an exponent $a > 1$, such that the probability that a patent's indegree is at least $k$ is proportional to $k^{-a}$. A smaller exponent indicates a heavier right tail. The degree distribution's power-law tail is a property of a standard preferential attachment model of network formation. Our model has two complication factors: the additional dimension of heterogenous quality $z$ among nodes and, more importantly, the depreciation of $z$ over time. We use the mean-field approximation approach to analyze the tail property of the stationary marginal distribution of indegrees. Detailed proofs are relegated to Appendix B.2.

**Proposition 2.** *(Power-law right tail of patent citation distribution.) Consider a sufficiently large size of the patent citation network as $t \to \infty$. For any fixed distribution of initial quality $Z$ over $[z_{\min}, z_{\max}]$, the mean-field approximation of the cross-sectional indegree distribution has a power-law right tail. The CCDF's tail exponent is*

$$\frac{1}{\bar{r}} = \frac{1}{(1-\bar{z})(1-\bar{\eta})}. \tag{16}$$

Proposition 2 shows which variables determine the shape of the stationary cross-sectional indegree distribution. In this model, the rate of exact citation replication $\bar{r}$ sets the pace at which widely cited patents attract disproportionally even more new citations. A high replication rate favors the "rich get richer" dynamics, resulting in higher probability mass at the right end of indegree distribution, or a heavier right tail. It happens when a new patent shares more similarities with its parent on average (low $\bar{z}$), and when the replication mutation rate ($\bar{\eta}$) is low.

### 4.3 Firm production decisions

Firms are risk neutral and discount future profits at the same rate $\rho > 0$ as the households. Conditional on the existence of a corresponding patent, the production technology is fully summarized by a simple production function, given as

$$y(\omega, t) = l(\omega, t), \quad \forall \omega \in \mathcal{N}(t),\ \forall t,$$

where $l(\omega, t)$ is the production labor input and $y(\omega, t)$ is the output level. A firm $f$ owns $n^f(t)$ patents at $t$ and hence operates as many product lines, each of which corresponds to a product variety. The firm takes as given the product-specific demand curve given in eq. (12) and the unit wage rate, and optimally prices each variety to maximize profits. Firms do not face any financial constraint. A standard optimal pricing rule follows, such that $p(\omega, t) = \frac{\nu}{\nu-1}$. The ideal price index $P(t)$ satisfies $P(t) = \frac{\nu}{\nu-1}\left(N(t)\overline{z^o}\right)^{-\frac{1}{\nu-1}}$. When the market for production labor clears at $\overline{L} = \int_{\omega\in\mathcal{N}(t)} l(\omega, t)$, we have $C(t) = \left(N(t)\overline{z^o}\right)^{\frac{1}{\nu-1}}\overline{L}$ and the aggregate flow profit resulting from production and sales $\pi \equiv \frac{1}{\nu-1}\overline{L}$. An equilibrium requires that $\Pi = \pi - w_R\overline{R}$ in the household's budget. The optimal output of each variety is then pinned down. The maximum instantaneous profit generated by a product line $\omega$ with current quality $z(\omega, t)$ satisfies

$$\pi(\omega, t) = \frac{\pi}{N(t)}\frac{z(\omega, t)}{\overline{z^o}}.$$

If a firm successfully develops a new patent $\omega$ and hence a new product line with initial quality $z$ at time $t$, then at $t+\tau$, $\omega$'s quality depreciates to $z(\omega, t+\tau) = ze^{-\delta\tau}$. Therefore, the present value of the additional profit stream brought by $\omega$ is given as

$$\int_0^\infty e^{-\rho\tau}\frac{\pi(\omega, t+\tau)}{N(t+\tau)}\mathrm{d}\tau = \frac{\pi}{N(t)}\frac{z(g+\delta)}{\overline{z}g(\rho+g+\delta)} \equiv \frac{z}{N(t)}\mathcal{P}(g, \overline{z}), \tag{17}$$

and $\mathcal{P}(g, \overline{z})$ represents the total production value of a new patent, when the aggregate knowledge grows at rate $g$ and the average initial patent quality is $\overline{z}$. It follows that $\partial_g\mathcal{P} < 0$ and $\partial_{\overline{z}}\mathcal{P} < 0$. The present value of a product line decreases in the growth rate $g$ of the total number of varieties and in the average initial quality $\overline{z}$ of a competing product. More varieties increase competition among firms, and so a newer variety's present value of profit also declines in time of birth. A higher quality $z$ increases the present value. Therefore, the positive present value of profit stream incentivizes firms to innovate and create high-quality new production lines.

### 4.4 Firm innovation decisions

The goal of this section is to examine how the citation network dynamics described in subsection 4.2 affect firms' innovation incentives and decisions. Firms utilize existing knowledge to innovate. At any time $t$, a firm $f$'s innovation decision involves three parts. First, firm $f$ decides whether to innovate in each sector of its technology space $j \in \mathcal{J}^f$. Second, conditional on innovating in sector $j$, firm $f$ observes its sector-specific innovation efficiency and decides its research intensity, which determines the firm's expected new knowledge (patents) output in $j$. Third, firm $f$ sets the target quality of each of its new patents in sector $j$. In what follows, we start from the second and third parts of a firm's innovation decision, and then we step back to the first part of the decision.

Consider firm $f$ that innovates in sector $j \in \mathcal{J}^f$ at time $t$. Firm $f$ has knowledge stock in $j$ given as $n_j^f(t) > 0$. The aggregate knowledge stock is $N(t)$. The firm draws an innovation efficiency $\epsilon_j^f(t)$ from an invariant distribution $G$ over $\mathbb{R}_+$. The draws are independent and identically distributed (i.i.d.) across firms and sectors, and the distribution $G$ has finite first and second moments. Given $\epsilon_j^f(t)$, a constant-returns-to-scale (CRS) Cobb-Douglas production function determines the expected increment of new knowledge at $t$, given as

$$\left(\epsilon_j^f(t) n_j^f(t)\right)^{1-\gamma} \left(\frac{\overline{N(t)} R_j^f(t)}{\gamma}\right)^{\gamma} \quad \text{s.t. } \gamma \in (0,1). \tag{18}$$

In the production function, $R_j^f(t)$ is the research labor input firm $f$ uses to innovate in sector $j$ at $t$, and $\overline{N(t)}$ is a compound of aggregate knowledge stock at $t$ and it is common to all firms, such that $\overline{N(t)} = \overline{A} N(t)$ for some $\overline{A} > 0$.[7] Let firm $f$'s *research intensity* in sector $j$ at time $t$ be $r_j^f(t)$, defined as the ratio of firm $j$'s research labor input over its knowledge stock,

$$r_j^f(t) \equiv \frac{R_j^f(t)}{n_j^f(t)}.$$

Then eq. (18) becomes the production function for $f$'s expected knowledge growth rate, given as

$$g_j^f(t) = \left(\epsilon_j^f(t)\right)^{1-\gamma} \left(\frac{\overline{A} N(t) r_j^f(t)}{\gamma}\right)^{\gamma}. \tag{19}$$

The production of $g_j^f(t)$ in eq. (19) is independent of the growth history. Given the aggregate knowledge compound $\overline{A}N(t)$ and the realization of innovation efficiency $\epsilon_j^f(t)$, the cost minimization problem $\min w_R r_j^f(t)$ to achieve a knowledge growth rate of $g_j^f(t)$ yields the following cost function for the minimal research intensity,

$$\frac{1}{N(t)} w_R c_r \left(g_j^f(t), \epsilon_j^f(t)\right) = \frac{1}{N(t)} w_R \frac{\gamma}{\overline{A}} \left(g_j^f(t)\right)^{\frac{1}{\gamma}} \left(\epsilon_j^f(t)\right)^{-\frac{1-\gamma}{\gamma}}. \tag{20}$$

Conditional on innovating in sector $j$ at time $t$, firm $f$'s decision of research intensity and knowledge output boils down to the optimal choice of its knowledge growth rate $g_j^f(t)$.

New knowledge is created as new patents. As described in subsection 4.2, a newborn patent cites $\overline{x}$ existing patents on average. Establishing a citation edge involves a payment by the citing firm to the cited firm. Let the per-citation payment be $\phi(t) = \frac{\phi}{N(t)}$.[8] Citation-related payments occur only when new patents are born. For example, when firm $f_1$ develops a new patent at $t$ and that new patent cites one of the patents owned by firm $f_2$, then $f_1$ pays $\phi(t)$ to $f_2$ at $t$. Therefore,

[7] The linear representation of the knowledge compound is with little loss of generality. When $N_j(t)$ in each sector $j$ grows at the common rate $g$, a linear aggregation is equivalent to any CRS aggregation of knowledge stocks by sector, in which case $\overline{A}$ summarizes the composition of knowledge stock by sector.

[8] Parameter $\phi > 0$ is exogenous because it is a per-citation payment. Mechanically, when a new patent cites existing ones, the number of outgoing edges from the new one always equals the total number of inward edges received by existing ones, regardless of $\phi$. However, as discussed next, patent quality is an endogenous choice and it affects the expected payment stream attracted by a new payment. Therefore, across firms or across sectors, the payment flows are endogenous as well.

every new patent requires a total of $\frac{1}{N(t)}\phi\overline{x}$ expected citation payment.

We emphasize that our notion of citation payment $\phi(t)$ is not literal. In a reduced form, it captures the present value of any payment from the citing firm $f_1$ to the cited firm $f_2$ when $f_1$ finds new application of $f_2$'s knowledge. For example, $f_1$ may start buying $f_2$'s products or renting $f_2$'s (tangible or intellectual) capital to use as inputs.

Conditional on innovation, firms decide each new patent's *target* initial quality $z \in [z_{\min}, z_{\max}]$. We focus on the interior choice of $z$. The target initial quality chosen for a patent serves as the *mean* of the quality distribution from which this patent randomly draws its actual initial quality.[9]

The cost of choosing a high target $z$ is summarized by a strictly convex cost function $c_z : \mathbb{R}_+ \to \mathbb{R}_+$, such that a higher-quality new patent requires increasingly more research labor, i.e., $c_z' > 0$ and $c_z'' > 0$. The additional research labor required by a new patent born at time $t$ with quality $z$ is $\frac{1}{N(t)} c_z(z\overline{x})$. The cost function also reflects the idea that quality $z$ corresponds to the originality of a patent compared with its parent patent, which in probability is the most similar existing patent to the new one. Developing original content that differs from the parent patent involves making original citations ($z\overline{x}$) independently, and that is increasingly difficult.

One benefit of having a high target quality $z$ is the high present value of profit stream generated by the production line, captured in eq. (17). Another benefit of having a high target quality $z$ depends on expected *future* network dynamics, described in subsection 4.2. A patent of high initial quality is expected to attract more citations in the future and brings more payments in present value. Therefore, the law of motion of a patent's indegree is payoff relevant and it depends on the choice of $z$. Firms know the depreciation rate $\delta$ and the mutation rate $\overline{\eta}$. Each innovating firm also takes as given other firms' innovation decisions summarized by a constant expected network growth rate $g$ and the constant average quality $\overline{z}$. The expected number of citations a patent with target initial quality $z$ receives at age $\tau$ is simply $k(z, \tau)$ given in Proposition 1, as $k(\cdot, \tau)$ is linear in the *actual* initial quality. The present value of the gross benefit to set the target initial quality of a new patent at time $t$ to be $z$ therefore satisfies

$$\int_0^\infty e^{-\rho\tau} \frac{\phi}{N(t+\tau)} \partial_\tau k(z, \tau) \mathrm{d}\tau \equiv z \frac{\phi\overline{x}}{N(t)} \mathcal{D}(g, \overline{z}),$$

where $\partial_\tau k(z, \tau)$ is the derivative of eq. (15) with respect to age $\tau$. Function $\mathcal{D}(\cdot)$ captures the benefit of patent quality due to degree growth. It is the rate at which a patent expects to gain more citations as initial quality increases, adjusting for time discount and network expansion. It is a feature of the network formation process, driven by each patent's expected degree dynamics, and common to all firms. In equilibrium, the collective (future) decisions of all firms pin down $\mathcal{D}(g, \overline{z})$

[9] More accurately, firms choose a quality distribution $Z$ from the set of all probability distributions with support $[z_{\min}, z_{\max}]$. As only the mean $z$ of any chosen distribution is payoff relevant, we focus on the choice of $z$ directly and leave the exact form of $Z$ unspecified. For example, $Z$ can be degenerate, and then the chosen mean $z$ is the actual initial quality.

through $g$ and $\overline{z}$. Using the definition of the exact replication rate $\overline{r} = (1-\overline{z})(1-\overline{\eta})$, we have

$$\mathcal{D}(g,\overline{z}) = \left(\frac{\overline{\eta}}{\overline{z}} + 1 - \overline{\eta}\right) \frac{(g+\delta)(\rho+g)}{(\rho+g+\delta)[\rho+g\overline{\eta}+g\overline{z}(1-\overline{\eta})]}. \tag{21}$$

Observe that as a function of the aggregate growth rate $g$ and the average initial quality $\overline{z}$, $\mathcal{D}(g,\overline{z})$ increases in $g$ and decreases in $\overline{z}$. Intuitively, these derivatives reflect two opposing forces that determine innovation benefit due to degree growth. One is the *innovation complementarity* effect, captured by $\partial_g \mathcal{D} > 0$. A firm benefits from faster innovation of other firms because its existing patents can attract more citation payments. Given $\overline{z}$, the innovation complementarity effect determines the upper and lower bounds of citation payment flows, captured by $\mathcal{D} \in \left(\frac{1}{\overline{z}}\frac{1-\overline{r}}{1+\rho/\delta}, \frac{1}{\overline{z}}\right)$. The other is a crowding-out effect, related to *creative destruction*, captured by $\partial_{\overline{z}} \mathcal{D} < 0$. The relative quality of a firm's patent declines when other firms choose high quality, especially with knowledge obsolescence, as younger patents' qualities experience less depreciation, which becomes a negative impact on older patents' ability to attract new citations.

Combining all costs and benefits associated with patent creation, a firm chooses a patent's target quality $z^*(g,\overline{z},w_R)$ to maximize the expected net benefit of each new patent. The optimization problem can be written as

$$\mathcal{B}(g,\overline{z},w_R) \equiv \max_{z\in[z_{\min},z_{\max}]} \Big\{ \overbrace{z\mathcal{P}(g,\overline{z})}^{\text{production value}} + \overbrace{z\phi\overline{x}\mathcal{D}(g,\overline{z})}^{\text{degree value}} - \underbrace{w_R c_z(z\overline{x})}_{\text{quality cost}} \Big\} - \underbrace{\phi\overline{x}}_{\text{payments}}, \tag{22}$$

where $\mathcal{P}(\cdot)$ and $\mathcal{D}(\cdot)$ are given in eqs. (17) and (21), respectively. An interior solution $z^*(\cdot)$ for the optimal target quality must satisfy

$$z^*(g,\overline{z},w_R)\overline{x} = \left(c_z'\right)^{-1}\left(\frac{1}{w_R\overline{x}}\left(\mathcal{P}(g,\overline{z}) + \phi\overline{x}\mathcal{D}(g,\overline{z})\right)\right). \tag{23}$$

A high research wage or a high average initial quality reduces the optimal target quality of any innovating firm, captured by $\partial_{w_R} z^* < 0$ and $\partial_{\overline{z}} z^* < 0$, respectively. The response in the optimal target quality to changes in the aggregate knowledge growth rate reflected by the sign of $\partial_g z^*$ depends on whether the innovation complementarity effect ($\partial_g \mathcal{D} > 0$) offsets product market competition ($\partial_g \mathcal{P} < 0$).

The optimal choice of each patent's target quality pins down the maximum present value of the net (expected) benefit $\mathcal{B}(g,\overline{z},w_R)$ of every new patent, given in eq. (22). Then $\frac{1}{N(t)}\mathcal{B}(g,\overline{z},w_R)$ is the maximum net benefit of a new patent born at $t$. By the envelope theorem, we see that the benefit decreases in research wage and average initial quality captured by $\partial_{w_R}\mathcal{B} < 0$ and $\partial_{\overline{z}}\mathcal{B} < 0$, and that the effect of knowledge growth $\partial_g \mathcal{B}$ remains ambiguous.

We are ready to characterize the optimal choice of a firm $f$'s innovation rate $g_j^f(t)$ in sector $j$, taking as given other agents' decisions. Anticipating the optimal choice of target quality, firm $f$ gets a flow innovation benefit of $\frac{1}{N(t)}\mathcal{B}(g,\overline{z},w_R)g_j^f(t)$ if its innovation rate in sector $j$ is $g_j^f(t)$.

The flow cost function $c_r(\cdot)$ associated with $g_j^f(t)$ after observing the innovation efficiency $\epsilon_j^f(t)$ is given in eq. (20). Firm $f$'s optimal choice of $g_j^f(t)$ conditional on innovating in $j$ solves

$$\max_{g_j^f(t)} \mathcal{B}(g, \bar{z}, w_R) g_j^f(t) - w_R c_r \left( g_j^f(t), \epsilon_j^f(t) \right). \tag{24}$$

We focus on the parameter subspace such that $\mathcal{B}(g, \bar{z}, w_R) > 0$. The following result is immediate.

**Proposition 3.** *Consider firm $f$ that innovates in sector $j$ at time $t$, taking as given a constant expected knowledge growth rate $g$ in each sector, the constant average initial quality $\bar{z}$, and the market wage rate $w_R$ for research labor. The current share of firm $f$'s sector-$j$ knowledge is $s_j^f(t) \equiv \frac{n_j^f(t)}{N(t)} = s$. The realization of its innovation efficiency is at $\epsilon_j^f(t) = \epsilon$. Conditional on innovating in sector $j$, firm $f$'s optimal target quality of any new patents in sector $j$ is $z^*(g, \bar{z}, w_R)$ given in eq. (23); and firm $f$'s optimal choice of its knowledge growth rate in sector $j$ solves eq. (24), given as $g^*(\epsilon;\ g, \bar{z}, w_R) = \epsilon \mathcal{G}(g, \bar{z}, w_R)$, where*

$$\mathcal{G}(g, \bar{z}, w_R) \equiv \left( \frac{\overline{A} \mathcal{B}(g, \bar{z}, w_R)}{w_R} \right)^{\frac{\gamma}{1-\gamma}}, \tag{25}$$

*where the net benefit of firm innovation $\mathcal{B}(\cdot)$ is given in eq. (22). The maximum flow payoff to firm $f$ for innovating in sector $j$ at time $t$ is $s\epsilon\mathcal{V}(g, \bar{z}, w_R)$, with*

$$\mathcal{V}(g, \bar{z}, w_R) \equiv (1-\gamma) \left( \frac{\overline{A}}{w_R} \right)^{\frac{\gamma}{1-\gamma}} \left( \mathcal{B}(g, \bar{z}, w_R) \right)^{\frac{1}{1-\gamma}}, \tag{26}$$

*where $\epsilon\mathcal{V}(\cdot)$ is the maximum of eq. (24).*

A feature of our model is that innovating firms make decisions along the *intensive margins* (quality choice) of the *extensive margins* (new patent growth rates). Both decisions depend non-trivially on the endogenous dynamics of the citation network summarized by $\mathcal{D}(g, \bar{z})$.

Lastly, we step back to the first part of a firm's innovation decision, taking as given the optimal choices of knowledge growth and patent quality. Specifically, firms decide whether to innovate in each sector by weighing the present values of their options.

The option value of not innovating in a sector is normalized at zero. Every firm innovating in a sector incurs a lump sum cost in research labor $\kappa(t) = \frac{1}{N(t)}\kappa$, $\kappa > 0$, for each patent it currently owns in that sector. At time $t$, if firm $f$ with knowledge share $s_j^f(t)$ decides to innovate in $j$ after observing the innovation efficiency $\epsilon_j^f(t)$, it must pay $s_j^f(t) w_R \kappa$ upfront. Therefore, firm $f$ finds it optimal to innovate in sector $j$ if and only if $s_j^f(t)\epsilon_j^f(t)\mathcal{V}(g, \bar{z}, w_R) - s_j^f(t) w_R \kappa \geq 0$. A firm's decision to innovate in a sector is hence a cutoff rule characterized by a threshold of innovation efficiency, which is independent of firm or sector characteristics.

**Proposition 4.** *Given $g$, $z$, and $w_R$, firm $f$'s optimal innovation decision is a cutoff rule characterized by*

*$\epsilon^*(g, \bar{z}, w_R)$, such that*

$$\epsilon^*(g, \bar{z}, w_R) = \frac{w_R \kappa}{\mathcal{V}(g, \bar{z}, w_R)}, \tag{27}$$

*and firm $f$ innovates in sector $j$ at $t$ iff its innovation efficiency is sufficiently high, $\epsilon_j^f(t) \geq \epsilon^*(g, \bar{z}, w_R)$.*

Proposition 4 is consistent with Fact 2. Firms with larger technology spaces are more likely to draw at least one innovation efficiency above the threshold, and then produce patents with high realized quality. Therefore, these firms are bigger and more likely to become pathfinders. As expected, the threshold efficiency increases in research wage and average initial quality, and the direction of change is unclear in knowledge growth rate.

### 4.5 Equilibrium

So far, we take as given the constant expected knowledge growth rate $g$, the average initial quality $\bar{z}$, and the research wage rate $w_R$ when analyzing firm decisions. To close the model with of rational expectations, the tuple $(g, \bar{z}, w_R)$ must be consistent with firm decisions. The derivation details are in Appendix B.5.

Rational expectations require the expected aggregate knowledge growth rate $g$ that firms take as given to be consistent with their optimal innovation decisions. Therefore, such a $g$ must satisfy

$$g = \mathcal{G}(g, \bar{z}, w_R)[1 - G(\epsilon^*)] \tag{28}$$

where $\mathcal{G}(\cdot)$ is given in eq. (25) and the cutoff efficiency $\epsilon^* = \epsilon^*(g, \bar{z}, w_R)$ is defined in eq. (27). The expected knowledge growth rate in each sector is also $g$, and so is the actual sectoral knowledge growth rate, by the law of large numbers.

Quality choice is independent of a firm's characteristics.[10] It follows that when all firms anticipate the same $(g, \bar{z}, w_R)$ in equilibrium, they pick the same target initial quality for each new patent. Therefore, $\bar{z}$ as the average quality of newborn patents must satisfy

$$\bar{z} = z^*(g, \bar{z}, w_R), \tag{29}$$

where $z^*(\cdot)$ is defined in eq. (23) as a firm's optimal choice of target quality for each of its new patents.

Lastly, the market price $w_R$ clears the market for research labor. The supply is fixed at $\overline{R}$. Conditional on innovation, a firm's demand for research labor consists of three parts — the fixed innovation cost, the research labor input to produce new patents, and the research labor required to achieve the target patent quality. Applying a law of large numbers yields the market clearing

[10] It is straightforward to generalize the model to allow for *endogenous* heterogeneity in the optimal choice of target quality and hence differing optimal innovation rates. Such a direct extension helps explain Fact 3. The simplest way is to allow the cost function $c_z(\cdot)$ to be firm or firm-sector specific, such that the marginal cost $c_z'(\cdot)$ also varies across firms or firm-sector pairs. The quality distribution $Z$ then becomes endogenous. Moreover, there can be persistent difference in patent quality and innovation rate across firms. The rest of the model remains intact. We focus on the baseline model to simplify the equilibrium conditions.

condition as follows,

$$\overline{R} = [1 - G(\epsilon^*)]\Bigg( \underbrace{\kappa}_{\substack{\text{fixed}\\\text{cost}}} + \underbrace{\frac{\gamma \mathcal{V}(g, \overline{z}, w_R)}{(1-\gamma) w_R}}_{\text{cost of quantity}} + \underbrace{\mathcal{G}(g, \overline{z}, w_R) c_z(\overline{zx})}_{\text{cost of quality}} \Bigg), \tag{30}$$

where $\mathcal{V}(\cdot)$ is given in eq. (26) and $\mathcal{G}(\cdot)$ in eq. (25).

The three equations eqs. (28), (29) and (30) generate the tuple $(g, \overline{z}, w_R)$, which in turn determines the rest of the equilibrium outcome. Therefore, we can directly define an equilibrium as a fixed point in the $(g, \overline{z}, w_R)$-space.

**Definition 1.** A *stationary rational-expectations equilibrium* is a tuple $(g, \overline{z}, w_R)$ with $g, w_R > 0$ and $\overline{z} \in [z_{\min}, z_{\max}]$ such that $(g, \overline{z}, w_R)$ solves the system of eqs. (28), (29) and (30). Equilibrium outcome is captured by a cutoff efficiency $\epsilon^* = \epsilon^*(g, \overline{z}, w_R)$, innovation payoff $\mathcal{V} = \mathcal{V}(g, \overline{z}, w_R)$, production value $\mathcal{P} = \mathcal{P}(g, \overline{z})$, and degree value $\mathcal{D} = \mathcal{D}(g, \overline{z})$, where the functions $\epsilon^*(\cdot)$, $\mathcal{V}(\cdot)$, $\mathcal{P}(\cdot)$, and $\mathcal{D}(\cdot)$ are given in eqs. (27), (26), (17) and (21), respectively.

### 4.6 Patent and firm values

Before we write down a firm's value in equilibrium at a given time, it is helpful to discuss the value of an existing patent and the associated production line. A firm's value then depends on the total value of all the patents it owns. The details of derivation are in Appendix B.6.

In equilibrium, the unconditional expectation of the time-$t$ present value of a *new* patent born at that time with initial quality $z$ is fully summarized by $\frac{1}{N(t)}(z\mathcal{P} + z\phi\overline{x}\mathcal{D})$, after the owner firm pays the relevant costs but before any quality depreciation and before attracting any edges from newer patents. As a patent ages, its value evolves, as its quality depreciates and its indegree increases. Consider an existing patent at time $t$ with *current* quality $\widetilde{z}$ and *current* indegree $d$. The time-$t$ present value of the associated product line is

$$\frac{1}{N(t)} \widetilde{z}\, \mathcal{P}$$

by eq. (17). The law of motion of the patent's expected indegrees at $t + \tau$ can be described by an identical differential equation to eq. (14), but with the initial condition given as $k(\widetilde{z}, 0) = d$. It follows that the time-$t$ present value of the expected citation payments attracted to this patent is

$$\frac{1}{N(t)} \left( \widetilde{z}\, \phi \overline{x} \mathcal{D} + d\phi \frac{g\overline{r}}{\rho + g(1-\overline{r})} \right),$$

where $\overline{r} = (1-\overline{\eta})(1-\overline{z})$ is the exact replication rate in equilibrium. The first part of the value is driven by the current quality $\widetilde{z}$ of the existing patent, which declines as it ages; whereas the second part increases in the patent's number of accumulated citations, which is expected to increase as it ages. Combining the two parts yields the time-$t$ present value $\frac{1}{N(t)} v(\widetilde{z}, d)$ of an existing patent with current quality $\widetilde{z}$ and current indegree $d$, such that

$$v(\widetilde{z},d) = \underbrace{\widetilde{z}\,\mathcal{P}}_{\substack{\text{remaining}\\ \text{production value}}} + \overbrace{\phi\overline{x}\left(\widetilde{z}\,\mathcal{D} + \frac{d}{\overline{x}}\frac{g\overline{r}}{\rho+g(1-\overline{r})}\right)}^{\text{remaining degree value}}. \tag{31}$$

The present value of *all* existing patents in the economy at any time is constant at $v(\overline{z^o},\overline{x})$, by the law of large numbers. Its expression reduces to

$$v(\overline{z^o},\overline{x}) = \frac{1}{\rho+g+\delta}\left(\pi + g\phi\overline{x}\frac{\rho+g+\delta\overline{r}}{\rho+g(1-\overline{r})}\right).$$

The present value of a firm $f$ at any time $t$ consists of three parts: profit streams generated by existing product lines, payment streams due to expected new citations of existing patents, and expected net benefit of innovation. The firm's time-$t$ knowledge stock is responsible for the first two parts, and its innovation decision results in the third. The following proposition shows that, along the equilibrium path characterized by $(g,\overline{z},w_R)$, the time-$t$ value of any firm $f \in \mathcal{F}$ is linear in the firm's knowledge stock share defined as $s^f(t) = \frac{n^f(t)}{N(t)}$. The knowledge share $s^f(t)$ is also the firm's output, sales, and employment share. The firm's value per share increases in the firm's average patent quality $\overline{z}^f(t)$ and the average number of accumulated citations per patent $\overline{d}^f(t)$.

**Proposition 5.** *(Equilibrium firm value.) Consider the path of a stationary equilibrium characterized by $(g,\overline{z},w_R)$ and the associated $\epsilon^*$, $\mathcal{V}$, $\mathcal{P}$, and $\mathcal{D}$. Suppose that, at time $t$, firm $f$ has a total market share of $s^f(t) = s$ with an average current patent quality $\overline{z}^f(t) = \widetilde{z}$ and an average number of accumulated citations per patent $\overline{d}^f(t) = d$. The time-$t$ value of firm $f$ is proportional to $s$, given as*

$$V(s,\widetilde{z},d) = s\left[\overbrace{\widetilde{z}\,\mathcal{P} + \phi\overline{x}\left(\widetilde{z}\,\mathcal{D} + \frac{d}{\overline{x}}\frac{g\overline{r}}{\rho+g(1-\overline{r})}\right)}^{\textit{average value of an existing patent}} + \overbrace{\frac{1}{\rho}\left(\mathcal{V}^* - [1-G(\epsilon^*)]w_R\kappa\right)}^{\textit{innovation value}}\right],$$

*where $\mathcal{V}^* = \mathcal{V}\int_{\epsilon^*}^{\infty}\epsilon \mathrm{d}G(\epsilon)$. Firm value $V$ is strictly increasing in $s$, $\widetilde{z}$, and $d$.*

Proposition 5 corresponds to Fact 4. The intuition is as follows. The value of each existing patent captured by eq. (31) is a linear combination of the patent's current quality and current indegree. Therefore, given firm $f$'s market or knowledge share $s^f(t)$, the within-firm mean of time-$t$ quality $\overline{z}^f(t)$ and that of accumulated indegrees $\overline{d}^f(t)$ are sufficient statistics to summarize existing patents' contributions to firm value. An implication is that a firm's value is history dependent, due to the randomness in both its innovation efficiency in each sector and every patent's actual number of accumulated citations. It is especially the case when a firm innovates in a small set of sectors. The last part of firm value is the expected net benefit of firm innovation, determined by firm decisions discussed in subsection 4.4. It is also proportional to a firm's current knowledge share as the share is expected to remain unchanged given the common expected knowledge growth rate $g$.

Proposition 5 establishes the connection between a firm's market value (an ex post quality measure) and the intrinsic quality of its patents. It further implies that a firm's market value is informative about the firm's average patent quality. In reality, a patent's intrinsic quality is difficult to measure. Jaffe's quality measures... and a theoretical support to that QJE... Our parsimonious way to model quality captures several aspects of what a patent's quality may be in reality, and how these aspects contribute to a firm's market value — the originality of the patent by definition of $z$, the direct product market value ($\tilde{z}\,\mathcal{P}$), the potential applicability ($\tilde{z}\,\phi\bar{x}\mathcal{D}$), and the acknowledged applicability ($d\phi\frac{g\bar{r}}{\rho+g(1-\bar{r})}$). Quantity ($s$) matters to a firm's value as is standard. Meanwhile, the size of a firm's technology space is the number of sectors in which the firm innovates, and so a large technology space helps diversify the risk of drawing a low innovation efficiency in each single sector.[11] Moreover, a larger technology space allows the firm more trials to get high quality realizations, increasing the chance of being a pathfinder.

Proposition 5 and eq. (31) also offer insights regarding whether and when the number of accumulated citations $d$ is predicative about a patent's and its owner firm's market value. ... *ceteris paribus*, $\phi$, $\bar{x}$, and age ($\tilde{z}$)

### 4.7 Firm growth and within-firm granularity

A stationary firm size distribution does not exist in our setup as firms never shrink in size, measured as the patent stock, and nor do they exhibit entry and exit dynamics. Nonetheless, the model sheds light on the mechanism of firm growth. Aggregation within a firm yields the equilibrium growth rate of a firm $f$ at any time $t$, which satisfies

$$g^f(t) = \sum_{j\in\mathcal{J}^f} \frac{n_j^f(t)}{n^f(t)} \mathbb{1}_{\{\epsilon_j^f(t)\geq\epsilon^*\}} \epsilon_j^f(t)\mathcal{G}.$$

In expectations, firm growth in our model is consistent with Gibrat's law. Regardless of a firm's size measured as the knowledge stock $n^f(t)$, its expected growth rate is always $g$, identical to the aggregate rate of knowledge growth.

In contrast, the variance of a firm's growth rate depends its characteristics; it is a function of the within-firm *Herfindahl index* across sectors in its technology space. Denote $H^f(t)$ as firm $f$'s Herfindahl index at time $t$, given as

$$H^f(t) = \sum_{j\in\mathcal{J}^f} \left(\frac{n_j^f(t)}{n^f(t)}\right)^2 \in \left[\frac{1}{J^f}, 1\right],$$

where $J^f \leq J$ is the number of sectors in firm $f$'s technology space $\mathcal{J}^f$. Then, as the variance of

[11] The argument for diversification is especially applicable if the opportunity to innovate arrives randomly for each firm and each sector, at a fixed Poisson rate. We can then reinterpret the fixed innovation cost $\kappa$ and the mean of the innovation efficiency's distribution $G$ such that they include the arrival rate of innovating opportunity.

innovation efficiency $(\sigma^*)^2 \equiv \text{Var}\left(\mathbb{1}_{\{\epsilon \geq \epsilon^*\}}\epsilon\right)$ exists and is finite, the variance of firm growth is

$$\text{Var}\left(g^f(t)\right) = (\mathcal{G}\sigma^*)^2 H^f(t). \tag{32}$$

A firm with large technology space and a more uniformly distributed knowledge stock has less volatility in growth as it can better diversify the risk associated with innovation efficiency. In contrast, if a firm is large with a large total stock of knowledge but it innovates in a small number of sectors, then its growth remains volatile. This prediction is confirmed by Fact 6.

Our model therefore provides a simple explanation to the slow decline in firm growth volatility as firms get larger, if any. It has been discussed by Klette and Kortum (2004); Luttmer (2011); Gabaix (2011), among others. The explanation is twofold. First, a firm's large knowledge stock within a sector does not reduce its growth volatility in that sector, in line with the standard Gibrat's law. Hence, knowledge growth rates of large but highly specialized firms remain volatile. Second, a firm's growth volatility may decrease slowly in the size of its technology space. Specifically, the decay speed is low when within-firm "granularity" exists, captured by a high Herfindahl $H^f(t)$. This within-firm granularity is analogous to the notion of granularity on the aggregate level by Gabaix (2011). When a firm's technology space is large, but the distribution of the firm's knowledge stock across sectors is highly skewed and concentrates at a few sectors, then the firm's knowledge growth in these sectors contributes disproportionally more to its overall growth. Growth volatility in these sectors translates to firm-level growth volatility and it is difficult to diversity within the firm.

### 4.8 Sector-level aggregation of the knowledge network

The network-formation process in this model is the continuous limit of a growing patent-level digraph with *unweighted* edges. New patents enter the network as a result of firm innovation. In this model, it is straightforward to group patents by sector and bundle the citation flows accordingly to produce an aggregated citation network. Aggregation results in a *weighted* digraph, with looping edges representing citations between patents that belong to the same sector. While the patent-level network has increasingly many vertices and edges, the aggregated network has a fixed number of vertices $J$ and growing counts of edges. Aggregation allows us to examine the model's implications at the sector level.

Consider two sectors $i, j \in \mathcal{J}$. Let $X^{i\to j}(t)$ be the accumulated number of citation edges going from sector-$i$ patents to sector-$j$ ones and $\dot{X}^{i\to j}(t)$ be the corresponding rate of accumulation. The total number of citation edges satisfy $X(t) \equiv \sum_{i,j\in\mathcal{J}} X^{i\to j}(t) = N(t)\overline{x}$ and $\dot{X}(t) \equiv \sum_{i,j\in\mathcal{J}} \dot{X}^{i\to j}(t) = gN(t)\overline{x}$. By the law of large numbers, each sector grows at the same constant rate $g$ and shares the same average current quality $\overline{z^o}$, and so each sector holds a constant share of patents $S_j = \frac{N_j(t)}{N(t)}$, $\forall j$ and $\forall t$. The law of motion of the number of citations going from $i$ to $j$ can be written as

$$\dot{X}^{i\to j}(t) = S_i g N(t) \left[\mathbb{1}_{\{i=j\}} + \overline{x}(1-\overline{z})\frac{X^{i\to j}(t)}{S_i N(t)\overline{x}} + (\overline{z}\overline{x} - 1)S_j\right]. \tag{33}$$

This equation is a sector-level counterpart to eq. (13) for a patent. When new nodes are born ($gN(t)$), a fraction $S_i$ of them are in sector $i$ and can potentially add $i$ to $j$ edges. The indicator function captures the child-parent links, which only happen within the same sector. Each sector-$i$ new patent is expected to replicate $\overline{x}(1-\overline{z})$ of its parent's citations, with and without mutations; and edges from $i$ to $j$ account for a fraction $\frac{X^{i\to j}(t)}{S_i N(t)\overline{x}}$ of all edges out of $i$. Note that the rate of mutation is irrelevant at the sector level, as mutation happens within the cited sector, which is $j$. Lastly, each new patent is expected to form $\overline{zx}-1$ quality-driven independent edges; nodes in $j$ attract such an edge with a total probability of $\frac{N_j(t)\overline{z^o}}{N(t)\overline{z^o}} = S_j$.

We examine the limiting behavior of cross-sector citations for tractability. Solving eq. (33) under appropriate initial conditions at $t=0$ yields a stationary distribution for cross-sector citations. It means that when there are sufficiently many firms innovating in each sector, the sector-level network can be well approximated using a constant weighted adjacency matrix $\mathbf{\Omega} = [\Omega_{ij}]_{J\times J}$, where a typical element is given as

$$\Omega_{ij} \equiv \frac{X^{i\to j}(t)}{X(t)} = S_i S_j \frac{\overline{zx}-1}{\overline{zx}} + S_i \frac{\mathbb{1}_{\{i=j\}}}{\overline{zx}}. \tag{34}$$

Note that, by definition, $\sum_{i,j\in\mathcal{J}} \Omega_{ij} = 1$ and $\dot{X}^{i\to j}(t) = \Omega_{ij}\dot{X}(t)$. Moreover, $\mathbf{\Omega}$ is symmetric, or $\Omega_{ij} = \Omega_{ji}, \forall i, j$.

Matrix $\mathbf{\Omega}$ offers a number of insights regarding sector-level innovations, despite its simplicity. Although the form of $\mathbf{\Omega}$ is derived as a stationary limiting case, the implications apply to more general cases where the ratios $X^{i\to j}(t)/X(t)$ and $\dot{X}^{i\to j}(t)/\dot{X}(t)$ are time-varying. First, sector-level knowledge stock and citation flows are tightly connected. Sectors with large knowledge stocks tend to attract as well as make more citations. Second, notice that we can interpret $\Omega_{ij}$ as the likelihood of a new edge going from $i$ to $j$. Between two sectors, the form of the likelihood resembles gravity, proportional to the product of patent numbers in both sectors. This aggregation result is consistent with our finding of Fact 5. Third, relatedly, $\Omega_{ij}$ captures the relative knowledge diffusion rate of new knowledge from sector $i$ to sector $j$. Knowledge diffusion between larger sectors or within a large sector is generally faster. Fourth, the extra terms along the diagonal of $\mathbf{\Omega}$ reflects a "home bias" that new patents tend to cite more existing ones in the same sector, consistent with the data. In this model, it is driven by the child-parent connections.

# 5 Quantitative Explorations

The previous section establishes that our theoretical model qualitatively captures the empirical findings. This section examines its quantitative performance. We begin by identification and calibration of key model parameters. Then, we discuss the implications of the empirically recovered measures of patent quality, its depreciation, and citation mutation rates. Lastly, we quantify the importance of network dynamics in firm innovation decisions and market values.

Table 1: Parameter Values

| Parameter | Value | Description |
|---|---|---|
| Panel A. Patent network formation | | |
| $\overline{x}$ | | Average citations made per patent |
| $\delta$ | | Patent quality depreciation rate |
| $\overline{\eta}$ | | Within sector mutation rate when copying parent's citation edges |
| Panel B. Households and firms: predetermined | | |
| $\rho$ | 0.03 | Discount rate |
| $\nu$ | 2.9 | Elasticity of substitution across product varieties (Broda and Weinstein, 2006) |
| $\overline{L}$ | 1 | Normalization of production-labor supply |
| $\overline{R}$ | 0.166 | Research-labor supply, 14.2% of total labor (Acemoglu et al., 2018) |
| $\overline{A}$ | 1 | Normalization of aggregate knowledge input |
| $\gamma$ | 0.5 | Innovation elasticity w.r.t. research input (Acemoglu et al., 2018) |
| Panel C. Households and firms: SMM | | |
| $\phi$ | | Citation payment |
| $\kappa$ | | Fixed cost of innovation |
| $\overline{c_z}$ | | Coefficient of quadratic quality cost $c_z(z\overline{x}) = \frac{\overline{c_z}}{2}(z\overline{x})^2$ |
| $G(\cdot)$ | | Distribution of innovation efficiency |

## 5.1 Identification and calibration

A simplifying feature of our model is that the network-formation part is separable from the rest of the equilibrium. The advantage is twofold. First, it enables us to calibrate the model parameters in separable groups. Specifically, we use patent citation data to directly identify and recover relevant parameters and variables, including average number of cites $\overline{x}$, quality depreciation rate $\delta$, average initial quality $\overline{z}$, and the exact replication rate $\overline{r}$ which then produces the value of mutation rate $\overline{\eta}$. Then, we assign values to the more standard parameters governing household preferences and firm production by following the literature. Lastly, we calibrate the rest of the parameters using simulated method of moments (SMM). Relatedly, the second advantage is that after we empirically recover values of $(\overline{x}, \delta, \overline{z}, \overline{r}, \overline{\eta})$ year by year, the theoretical model helps interpret any observed time trend or variation.

Table 1 reports the selected parameter values. In what follows, we discuss their identification and calibration in detail.

### 5.1.1 Network formation

$\overline{r}_t$

We identify the exact replication rate $\overline{r}$ using the prediction in Proposition 2 that citation distribution per parent follows a Pareto distribution with the tail shape parameter $\frac{1}{\overline{r}}$. We allow $\overline{r}$ to vary over time, and follow the method and code in Clauset, Shalizi, and Newman (2009). We find that $r = 1/3.5 = 0.2857$ consistently over time from 1976 to 2014. By definition, $r = (1-\overline{z})(1-\eta)$ is the share of new citations that exactly copy parent patent's citations. If $r$ keeps constant over time, then the exact copying citation share is static too.

$\overline{z}_t$

We estimate $\overline{z}_t$ using year specific version of (34).

$$\Omega_{ij,t} \equiv \frac{X_t^{i\to j}(t)}{X(t)} = S_{i,t}S_{j,t}\frac{\overline{z}_t\overline{x}_t - 1}{\overline{z}_t\overline{x}_t} + S_{i,t}\frac{\mathbb{1}_{\{i=j\}}}{\overline{z}_t\overline{x}_t}.$$

where $\overline{x}_t$ is the average number of citations per patent granted in year t in the USPTO citation data, we present $\overline{x}_t$ in the left panel of Figure 4. $S_{i,t}$ is sector i's patent stock share among all sectors in year t. In the following equation, we run OLS with restriction that $b_{1,t} + b_{2,t} = 1$ for every year t from 1976 to 2014.

$$\Omega_{ij,t} = b_{1,t} * S_{i,t}S_{j,t} + b_{2,t} * S_{i,t}\mathbb{1}_{\{i=j\}}.$$

Then $\overline{z}_t$ is estimated as

$$\overline{z}_t = \frac{1}{b_{2,t}\overline{x}_t}.$$

The estimated share of randomly attached citations $\overline{z}_t$ is finally presented in the right panel of Figure 4. We find that $\overline{z}_t$ is declining over time after 1976, or new patents are more likely to replicate their parents' citations in more recent years. $\overline{z}_t$ is also our measure of average patent quality, or the amount of knowledge spillovers one patent can contribute to others, especially to other patents that are further apart in the knowledge networks through random attachment. From consumer side, the decline of $\overline{z}_t$ also reduces welfare gain from a given rate of patent growth $g$.

By definition, if all patents in a sector i has never cited patents in sector j before year t, and suddenly a sector i patent cite a sector j patent in t+1, then this citation must be randomly attached, because there is no parental citations to copy from, and the model assume that random citations are more likely to happen among high quality patents. When average patent quality declines overtime, the model predicts that pathfinder patents, which cited a patent in a new technology category, become less likely. In Figure 5, we find supporting evidence that the share of pathfinder patents declines over time from 1976. Using published UCB Fung Institute Patent Data from Balsmeier, Assaf, Chesebro, Fierro, Johnson, Johnson, Li, Lück, O'Reagan, Yeh et al. (2018), we find similar phenomenon that novel word per patent has been declining since 1985, see Figure 6.

Our measure of cross-sector citation flow has a caveat, because the definition of technological classes has been stagnant after 1970s, very few new classes has been added to the US patent class, international patent classification (IPC) or collaborated patent classification (CPC). As a result, the cross-sector citation flow could be underestimated.

$\eta_t$

Equipped with $\overline{r}_t$ and $\overline{z}_t$, we can estimate $\eta_t$ as

$$\eta_t = 1 - \frac{\overline{r}_t}{1 - \overline{z}_t}$$

using Equation (16). The estimated $\overline{z}_t$ is finally presented in the middle panel of Figure 4. $\eta_t$ moves in the opposite direction of $\overline{z}_t$ over time. By definition, $(1 - \overline{z})\eta$ is the share of citations that follow a mutated parent patent's citation, meaning the citation is given to another patent within the same sector of the parent's cited patent. Since $\bar{z}_t$ decreases and $\eta_t$ increases, $(1-\bar{z})\eta$ must increase over time.

$\boldsymbol{\delta_t}$

The discrete-time version of eq. (14) is (allowing for time dependence)

$$k_{t+1}(\omega) - k_t(\omega) = g_{t+1}\frac{z(\omega, t+1)}{\overline{z^o_{t+1}}}(1-\bar{r}_{t+1})\bar{x}_{t+1} + g_{t+1}\bar{r}_{t+1}k_t(\omega),$$

where $z(\omega, t+1) = z(\omega)(1-\delta)^{\tau(\omega,t+1)} = z(\omega)\prod_{s=0}^{\tau(\omega,t)}(1-\delta_{t+s})$. Consider a cohort $t$ denoted $\Delta\mathcal{N}(t)$ at time $t+\tau$. The by-cohort aggregation becomes

$$\begin{aligned}
\Delta K^t_{t+\tau} = &\int_{\omega\in\Delta\mathcal{N}(t)}[k_{t+\tau}(\omega) - k_{t+\tau-1}(\omega)]\mathrm{d}\omega \\
=& g_{t+\tau}\frac{\int_{\omega\in\Delta\mathcal{N}(t)} z(\omega)\mathrm{d}\omega\prod_{s=1}^{\tau}(1-\delta_{t+\tau-s})}{\overline{z^o_{t+\tau}}}(1-\bar{r}_{t+\tau})\bar{x}_{t+\tau} + g_{t+\tau}\bar{r}_{t+\tau}\int_{\omega\in\Delta\mathcal{N}(t)} k_{t+\tau-1}(\omega) \\
=& g_{t+\tau}\frac{\Delta N_t\bar{z}_t\prod_{s=1}^{\tau}(1-\delta_{t+\tau-s})}{\overline{z^o_{t+\tau}}}(1-\bar{r}_{t+\tau})\bar{x}_{t+\tau} + g_{t+\tau}\bar{r}_{t+\tau}K^t_{t+\tau-1}
\end{aligned}$$

where $Z^t \equiv \Delta N_t\bar{z}_t$ is cohort-$t$'s aggregated initial quality at birth, $Z^t\prod_{s=1}^{\tau}(1-\delta_{t+\tau-s})$ is cohort-$t$'s quality at time $t+\tau$, and $\overline{z^o_{t+\tau}}$ the average current quality of *all* existing patents at the beginning of $t+\tau$. The issue is $Z^t$ and $\overline{Z^o}_{t+\tau}$. Observe that for two cohorts $t$ and $t+1$ in the same year $t+\tau$,

$$\frac{\Delta K^{t+1}_{t+\tau} - g_{t+\tau}\bar{r}_{t+\tau}K^{t+1}_{t+\tau-1}}{\Delta K^t_{t+\tau} - g_{t+\tau}\bar{r}_{t+\tau}K^t_{t+\tau-1}} = \frac{Z^{t+1}}{Z^t(1+\delta_t)} = \frac{\Delta N_{t+1}\bar{z}_{t+1}}{\Delta N_t\bar{z}_t(1-\delta_t)}$$

We use previously obtained $\bar{r}$ in LHS and $\bar{z}_t$ for RHS. Then for every year from 1976 to 1999, we obtain $\Delta K^t_{t+\tau}$ and $K^t_{t+\tau}$ for $\tau \in (1, 2, ..., 15)$ from the USPTO data. $\delta_t$ is then set to equate the RHS with the mean of 15 values of LHS with different citation lag $\tau$s. We do not find a significant time trend for $\delta_t$, therefore we assume $\delta$ is constant and equal to the average of 15 $\delta_t$ values, 0.1009.

#### 5.1.2 Households and firms

Several parameters in the model are common ingredients in the literature. We therefore set their values following standard practice. The unit length of time in the model is a year to match the data frequency. We set the discount rate at $\rho = 0.03$. We normalize the supply of production labor, i.e., the numeraire, to be one unit at $\bar{L} = 1$. We also normalize the efficiency $\bar{A} = 1$ in the aggregate knowledge compound as an input of new knowledge. The constant elasticity of substitution across product varieties in the consumption compound is set at a standard value $\nu = 2.9$, which is the median of the estimates by Broda and Weinstein (2006). We follow Acemoglu et al. (2018) in the choices of research-labor supply $\bar{R} = 0.166$ and its elasticity as innovation input $\gamma = 0.5$.

The rest of the parameters are calibrated by simulating the full model. Computation is relatively straightforward thanks to the analytical forms. By definition, the equilibrium outcome is fully captured by the fixed point $(g, \bar{z}, w_R)$ of the system of eqs. (28), (29) and (30). We can map $g$ and $\bar{z}$ directly to the patent citation data.

Parameter identification is as follows. The average probability that a firm innovates in a sector in its technology space pins down $1 - G(\epsilon^*)$, the only statistics of $G$ that enters the equilibrium condition. Combining eq. (28) and the expression for a firm's optimal knowledge growth in Proposition 3 yields $(g^*)^f_j(t) = \epsilon \mathcal{G} = \frac{\epsilon^f_j(t) g}{1-G(\epsilon^*)}$. We map the firm-sector specific knowledge growth $(g^*)^f_j(t)$ conditional on successful innovation to the data and recover the realized $\epsilon^f_j(t) \geq \epsilon^*$, $\forall f, j, t$. Then $\epsilon^*$ is set to be the minimum of these realizations, averaged across firm-sector pairs.

For simplicity, we assume a quadratic quality cost function such that $c_z(z\bar{x}) = \frac{\bar{c}_z}{2}(z\bar{x})^2$. Then we have three parameters left to be identified: the cost coefficient $\overline{c_z}$, the citation payment $\phi$, and the innovation fixed cost $\kappa$. Based on the set of calibrated parameters so far, we observe that these three parameters together with the equilibrium research wage $w_R$ can be viewed as the solution to a system of four equations and four unknowns that consists of the three equilibrium conditions eqs. (28), (29) and (30) and the cutoff efficiency eq. (27).

Before solving the four unknowns using four equations, we need to calibrate $G(\epsilon^*)$ and $\epsilon^*$. By definition, we estimate $1 - G_t(\epsilon^*)$ as the probability for firms to innovate in any sector j in time t. Since $\epsilon^f_j$ is independent across j and f, the probability to innovate in j is the same for incumbents and potential new entrants. Since we don't know how many potential new entrants are out there, we calculate innovation likelihood for sector j as the number of firms that innovate in sector j both in t and t-1 divided by the number of firms that innovate in sector j in t-1, which is the incumbent probability of innovation. Then we take average across all sectors and derive our estimation of $1 - G_t(\epsilon^*)$.

We back out $\epsilon^f_j(t)$ using $(\frac{1-G(\epsilon^*))}{g}(g^*)^f_t(t)$, and $(g^*)^f_t(t)$ is estimated as

$$(g^*)^f_t(t) = \frac{\text{number of new patents by firm f in sector j time t}}{\text{patent stock by firm f in sector j time t-1}} - 1.$$

Then we estimate $\epsilon^*_t$ as minimum level of all $\epsilon^f_j(t)$ in year t. We present the calibrated $G_t(\epsilon^*)$ and $\epsilon^*_t$ in Figure 7.

Given $G_t(\epsilon^*)$ and $\epsilon^*_t$, along with previously obtained $\eta_t$, $\bar{z}_t$, $\delta_t$, $\bar{x}_t$ and other parameters values, we finally can solve for $\phi_t$, $\kappa_t$, $\bar{c}_{z,t}$ and $w_R$ using eqs. (27), (28), (29) and (30). The parameters values are shown in Figure 8. We find that fixed cost of R&D $\kappa$ has been declining, however, marginal cost of R&D to increase patent quality $\bar{c}_z$ has been increasing especially quickly after 2000. $w_R$ has been slowly declining after 1980. The sharp increase in marginal cost of patent quality $\bar{c}_z$ explains the declining average patent quality $\bar{z}$ and likelihood to find a pathfinding patent since 2000, because only very high quality patents cite existing patents in other sectors without previous cross sector citation route. Higher marginal cost of research quality causes the aggregate productivity slow down since 2000.

### 5.2 Counterfactual Analysis

In this subsection, we look into the factors that contributed to aggregate productivity slow down since 1970s. We set the some of following parameter values $\overline{x}$, $\overline{c}_z$, $\eta$, $\phi$, $\kappa$, and $w_R$ at their 1976 initial values through the entire sample period one combination at a time, some endogenous variables other than $g$ will change along as well, such as $z^*$, $\epsilon^*$, $G(*)$ and $\overline{R}$, we use Equations eqs. (27), (28), (29) and (30) to pin down them in the counterfactual scenarios. To estimate $G_t(*)$, we assume $G_t(.)$ follows a lognormal distribution and estimate its mean and standard deviation using data implied $g_t^f$. Finally, we compare the model implied counterfactual time series of quality weighted growth rate $g\overline{z}$, which matters to consumer welfare, with their empirical counterparts.

In Figure 9, we find that setting average citation per patent $\overline{x}$ to 1976 level alone is the most effective to restore higher growth rate among single variable counterfactuals. Rising $\overline{x}$ over time (see the left panel of Figure 4) means to produce one new patents, inventors need more knowledge input in number of cited patents, either because the threshold to validate enough originality in a new patent is higher later on; or the amount of intrinsic knowledge embodied in each cited patent is smaller over time because of finer specialization; or inventors cite more other patents only to avoid potential law suit in the future. The fundamental reason could be lack of new sectors emerging every year, therefore inventions within existing sectors have to crowd out each other and firms put too much effort defending against each other.

In Figure 10, when we set 2 parameters back to their 1976 levels, the combination of citation per patent $\overline{x}$ and research wage $w_R$ is the best to maintain 1976's high growth rate. $\overline{x}$ is responsible for growth slow down after 1980s; $w_R$ is accountable for the growth slow down before 1980s, as we can see a rapid wage growth in the lower right panel of Figure 8. Otherwise, citation per patent $\overline{x}$ and marginal cost of patent quality $\overline{c}_z$ is the second best combination to push quality weighted growth rate back to 1980's level.

In Figure 10, among the 3-parameter combinations including $\overline{x}$ and $w_R$, $\overline{x}$, $w_R$ and $\overline{c}_z$ can almost perfectly replicate the 1976 growth rate through the whole sample period. Therefore, we conclude that the best policy combination is to reduce these three parameters back to their 1976 levels.

### 5.3 Policy Implications

The results of our calibration and counterfactual exercises all point to the rising cost components of R&D as factors that cause productivity growth slowdown since 1970s. First, IPR policy need to reduce the cost involved with citations to prior art, especially the number of citations per patent $\overline{x}$. Scientists need to be able to prove their patents' originality and novelty to the patent office or potential investors using signals other than extensive and costly citations to prior art. Inventors do not have to cite many previous patents only to pre-empt future IPR dispute. Ultimately, R&D and IPR policies need to encourage the creation of entire new sectors, which is the only solution to crowding out and over specialization within existing sectors. Otherwise the patent system itself may become the most important hinderance to productivity growth.

Second, current R&D policies have done a good job reducing fixed research cost as reflected in declining $\kappa$ in the upper right panel of Figure 8, but more effort need to be given to curb the marginal cost to improve patent quality $c_z$ (lower left panel of Figure 8). For example, high quality innovation tend to be riskier and more time consuming than low quality research, therefore subsidies targeting such high quality projects are more beneficial than universal subsidies to all type of research. Besides, our empirical evidence shows that, a special kind of high quality patents, pathfinder patents generate large positive externality to peer firms by paving new pathways in the knowledge networks that peer firms can follow in the future. R&D subsidies should target such pathfinding projects for their positive externalities.

Third, subsidy to R&D wage is still necessary. The well known skill biased technology change has increased high skill researchers' wage substantially. Tax breaks to high skilled workers doing R&D is one example to help firms reduce labor cost of research. Immigration policy that welcomes high skilled researchers also assures that there are adequate R&D workers in the labor market and keeps the wage rate down.

## 6 Conclusion

Knowledge networks matter at both micro and macro levels. We show that the creation of new citation links across sectors in knowledge networks increases firms innovation rate, market value, profit and productivity, not only for the firms that discovered the new citation links themselves, but also for other firms who innovate in these sectors. We then build a dynamic network formation model, where some citations are given to parental patent's citations or its related mutations within the same cited sector; other new citations are randomly assigned, and higher quality patents are more likely to give and attract random citations. The model implies that the expected quality for both the citing and cited patents of a new citation links across sectors must be exceptionally high, because such new path in the citation networks has no previous parental citation to follow. Peer firms in the same sector as cited patent benefit because they can may receive future mutated citations from offspring patents of the pathfinder patent; peer firms in the same sector as citing sector also gain because they can follow the newly discovered citation path in the future to find high quality knowledge input. After we calibrate the model using patent data, the counterfactual analysis reveals that rising R&D cost factors, such as number of backward citations per patent as knowledge input, marginal cost of patent quality, and researcher wage rate, explain the innovation rate slow down since 1976. R&D and IPR policies need to decrease the aforementioned cost factors in order to restore productivity growth.

## 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. 2017. "Endogenous production networks." Tech. rep., National Bureau of Economic Research.

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.

Bak, Per, Kan Chen, José Scheinkman, and Michael Woodford. 1993. "Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics." *Ricerche Economiche* 47 (1):3–30.

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 Rezza. 2018. "Cascading failures in production networks." *Econometrica* 86 (5):1819–1838.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb. 2020. "Are ideas getting harder to find?" *American Economic Review* 110 (4):1104–44.

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

Boldrin, Michele and David K Levine. 2013. "The case against patents." *Journal of Economic Perspectives* 27 (1):3–22.

Broda, Christian and David E Weinstein. 2006. "Globalization and the Gains from Variety." *The Quarterly journal of economics* 121 (2):541–585.

Cai, Jie and Nan Li. Forthcoming. "Growth through intersectoral knowledge linkages." *Review of Economic Studies* .

Cai, Jie, Nan Li, and Ana Maria Santacreu. 2019. "Knowledge Diffusion and Trade Across Countries and Sectors." *Federal Reserve Bank at St. Louis Working Paper* .

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

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

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

Dinopoulos, Elias and Constantinos Syropoulos. 2007. “Rent protection as a barrier to innovation and growth.” *Economic Theory* 32 (2):309–332.

Elliott, Matthew, Benjamin Golub, and Matthew O Jackson. 2014. “Financial networks and contagion.” *American Economic Review* 104 (10):3115–53.

Gabaix, Xavier. 2011. “The granular origins of aggregate fluctuations.” *Econometrica* 79 (3):733–772.

Grassi, Basile et al. 2017. “IO in IO: Size, industrial organization, and the input-output network make a firm structurally important.” Tech. rep.

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

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

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

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

Higham, Kyle, Gaétan de Rassenfosse, and Adam B Jaffe. 2020. “Patent Quality: Towards a Systematic Framework for Analysis and Measurement.” .

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.

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

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

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

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

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

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

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

König, Michael D. 2016. "The formation of networks with local spillovers and limited observability." *Theoretical Economics* 11 (3):813–863.

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

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

Luttmer, Erzo GJ. 2011. "On the mechanics of firm growth." *The Review of Economic Studies* 78 (3):1042–1068.

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

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

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

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

Taschereau-Dumouchel, Mathieu. 2017. "Cascades and fluctuations in an economy with an endogenous production network." .

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

Vázquez, Alexei. 2003. "Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations." *Physical Review E* 67 (5):056104.

# Appendix A Robustness Checks

In the regressions in (1) and (6), we treat every new cross-category citation links as equally important to the knowledge networks. Firms' exposure to knowledge network dynamics is weighted using firms' patent portfolio over different patent categories as specified in equations (2). In this section, we try different edge important measures as a second weight to calculate a more accurate measure of firm exposure to knowledge network dynamics.

First, we weight firm patent application number by inward citations they received as a measure of patent quality adjustment. The result is presented in Table C. We find that the result in Table **??** is preserved only for new links received. More own pathfinder patent even predicts lower inward citation weighted patent growth rate within 5 years, because patent inward citations as a patent quality measure take longer than 5 years to appear in the data.

Second, we calculate edge betweenness centrality (EBC) for annual knowledge networks, then we use EBC as weight to every new citation link, on top of firm's patent stock share, when calculating the double weighted number of pathfinder and path-receiver links to for each firm. The results presented in Tables 13 to 18 confirm the pattern we observe in Tables 5 to 10.

# Appendix B Proofs

## B.1 Proof of Proposition 1

Consider an arbitrary existing node $\omega \in \mathcal{N}(t)$. It has an industry classification $j(\omega)$, an initial quality $z(\omega)$, current age $\tau(\omega, t)$, and current indegree $d(\omega, t)$. We copy below eq. (13):

$$\frac{1}{N(t)}\left[\frac{z(\omega,t)}{\bar{z}g/(g+\delta)}\bar{x}(1-\bar{r})+d(\omega,t)\bar{r}\right],$$

where $z(\omega,t)=z(\omega)e^{-\delta\tau(\omega,t)}$ is $\omega$'s current quality. The network size $N(t)$ grows at rate $g$, then in a very short time period between $t$ and $t+\mathrm{d}t$, the network gets $gN(t)\mathrm{d}t$ of new nodes, and hence $\omega$'s indegree is expected to increase by

$$\mathbb{E}[d(\omega,t+\mathrm{d}t)\mid d(\omega,t)]-d(\omega,t)=\left[\frac{z(\omega,t)}{\bar{z}/(g+\delta)}\bar{x}(1-\bar{r})+d(\omega,t)g\bar{r}\right]\mathrm{d}t, \tag{B.1}$$

where $\mathbb{E}[\cdot\mid d(\omega,t)]$ is the conditional expectation operator, given the realized $d(\omega,t)$ at time $t$. By definition, $k(z,\tau)$ is the unconditional expectation of a node's indegree at age $\tau$ if it has an initial quality $z$. It is straightforward to see that, thanks to linearity, eq. (14) is the same expression above for the expected indegree after taking $\mathrm{d}t \to 0$. Then, solving eq. (14) using the boundary condition $k(z,0)=0$ produces eq. (15). Taking the first-order derivative of eq. (15) yields

$$\frac{\partial k(z,\tau)}{\partial\tau}=ze^{-\delta\tau}\frac{\bar{x}(1-\bar{r})(g+\delta)}{\bar{z}(\delta+g\bar{r})}\left[g\bar{r}e^{(\delta+g\bar{r})\tau}+\delta\right]. \tag{B.2}$$

### B.2 Proof of Proposition 2

We use the mean-field approximation approach to look for the stationary cross-sectional indegree distribution when the network size is sufficiently large. Following standard practice, we examine the distribution implied by the law of motion of expected indegrees in eq. (15), exploiting the large-sample properties when $N(t) \to \infty$ as $t \to \infty$.

The goal is to characterize the tail property of the limiting cross-sectional indegree distribution, or the fraction of nodes with more than $d$ citations denoted as $T(d)$, when $d$ is large. eq. (15) can be represented as

$$k(z,\tau) = z \times \xi(\tau),$$

where $\xi(\cdot)$ is strictly increasing in $\tau$ and so its inverse $\xi^{-1}(\cdot)$ is well-defined. The constant network growth rate $g$ implies an exponential stationary age distribution, with a CCDF given as $e^{-g\tau}$. Therefore, for each $z \in [z_{\min}, z_{\max}]$,

$$\Pr(k(z,\tau) > d) = \Pr\left(\tau > \xi^{-1}(d/z)\right) = e^{-g\xi^{-1}(d/z)}.$$

With a given distribution $Z$ of initial quality, the tail probability $T(d) = \mathbb{E}^z\left[e^{-g\xi^{-1}(d/z)}\right]$ represents the fraction of nodes with more than $d$ citations. It suffices to show that $T(d) = \mathbb{E}^z\left[e^{-g\xi^{-1}(d/z)}\right]$ approaches the right tail of a Pareto distribution when $d$ is large.

It remains to be shown the functional form of $\xi^{-1}(d/z)$. To ease expressions, define

$$\bar{\xi} = \frac{\bar{x}(1-\bar{r})(g+\delta)}{\bar{z}(\delta+g\bar{r})}.$$

Then we set

$$\xi(\tau) = \bar{\xi}e^{-\delta\tau}\left[e^{(\delta+g\bar{r})\tau} - 1\right] = \frac{d}{z} \quad \Longrightarrow \quad e^{g\bar{r}\tau} = \frac{d}{z\bar{\xi}} + e^{-\delta\tau}$$

and we solve for $\tau$. When $\delta > 0$, the equation does not have a closed-form solution. However, when $d$ is large and hence $\tau$ is large, we can take the logarithm of both sides and approximate the solution by

$$\tau = \xi^{-1}(d/z) \approx \frac{1}{g\bar{r}}\log\left(\frac{d}{z\bar{\xi}}\right).$$

Therefore, when $d$ is large, the tail probability $T(d)$ of the indegree distribution approaches the power law,

$$T(d) \approx \mathbb{E}^z\left[e^{-g\xi^{-1}(d/z)}\right] = \mathbb{E}^z\left[\left(z\bar{\xi}\right)^{\frac{1}{\bar{r}}} d^{-\frac{1}{\bar{r}}}\right] \sim d^{-\frac{1}{\bar{r}}}, \quad \text{as } d \to \infty.$$

This result concludes the proof.

Note that, at the other extreme when $d \to 0$ and hence $\tau \to 0$, we have $e^{g\bar{r}\tau} \approx 1 + (g\bar{r})\tau$ and

$e^{-\delta\tau} \approx 1 - \delta\tau$. The approximated solution to $\xi(\tau) = d/z$ becomes

$$\tau = \xi^{-1}(d/z) \approx \frac{d}{z\bar{\xi}(g\bar{r} + \delta)}.$$

It follows that the mean-field approximation of indegree distribution when $d \to 0$ is approximately exponential, and the approximated tail probability is

$$T(d) \approx \mathbb{E}^z \left[ e^{-g\xi^{-1}(d/z)} \right] = \mathbb{E}^z \left[ \left( e^{1/[z\bar{\xi}(\bar{r}+\delta/g)]} \right)^{-d} \right] \sim e^{-d/[\bar{z}\bar{\xi}(\bar{r}+\delta/g)]}, \quad \text{as } d \to 0.$$

### B.3 Proof of Proposition 3

The FOC of eq. (24) yields

$$\mathcal{B}(g, \bar{z}, w_R) = w_R \frac{\partial c_r \left( g_j^f(t), \epsilon_j^f(t) \right)}{\partial g_j^f(t)} = \frac{w_R}{\bar{A}} \left( g_j^f(t) \right)^{\frac{1-\gamma}{\gamma}} \left( \epsilon_j^f(t) \right)^{-\frac{1-\gamma}{\gamma}}.$$

The rest of Proposition 3 follows the assumed functional form of $\mathcal{V}(g, \bar{z}, w_R)$. Note that $\epsilon\mathcal{V}$ is the maximum of eq. (24), whereas a firm's flow benefit of innovation in the sector is $\epsilon\mathcal{V}$ multiplied by the firm's current share of knowledge stock in this sector $s$. It is due to the definition of firm-sector specific knowledge growth rate and the necessary $N(t)$ adjustment for present values.

### B.4 Proof of Proposition 4

The result is immediate as the cutoff innovation efficiency $\epsilon^*(g, \bar{z}, w_R)$ must satisfy

$$\epsilon^*(g, \bar{z}, w_R)\mathcal{V}(g, \bar{z}, w_R) = w_R\kappa,$$

where the functional form of $\mathcal{V}(\cdot)$ is in eq. (26), which yields the cutoff in eq. (27).

### B.5 Details of subsection 4.5: equilibrium conditions

In each industry $j$, the expected knowledge growth rate $g_j(t)$ is the expected weighted average of all innovating firms' optimal growth rates, given as

$$\begin{aligned} g_j(t) =& \mathbb{E} \left[ \frac{1}{N_j(t)} \int_{f \in \mathcal{F}} n_j^f(t) g^* \left( \epsilon_j^f(t);\ g, \bar{z}, w_R \right) \mathbb{1}_{\left\{ \epsilon_j^f(t) \geq \epsilon^*(g, \bar{z}, w_R) \right\}} \right] \\ =& \mathcal{G}\left( g, \bar{z}, w_R \right) \int_{f \in \mathcal{F}} \frac{n_j^f(t)}{N_j(t)} \mathbb{E} \left[ \epsilon_j^f(t) \mathbb{1}_{\left\{ \epsilon_j^f(t) \geq \epsilon^*(g, \bar{z}, w_R) \right\}} \right]. \end{aligned}$$

Recall that $N_j(t) = \int_{f \in \mathcal{F}} n_j^f(t)$ and that each $\epsilon_j^f(t)$ is an i.i.d. draw from $G$. Therefore, all industries share a common and time-invariant expected growth rate $g_j(t) = \mathcal{G}\left( g, \bar{z}, w_R \right) \mathbb{E} \left[ \epsilon \mathbb{1}_{\{\epsilon \geq \epsilon^*(g, \bar{z}, w_R)\}} \right] =$

$g$, which is eq. (28). The condition for $\bar{z}$ in eq. (29) is obvious.

Firms' demand for research labor comes from three sources. The first is the proportional fixed cost of innovation $\frac{n_j^f(t)}{N(t)}\kappa$ if the firm decides to innovate in the sector at $t$. The second is the extensive-margin research labor for knowledge growth conditional on innovation, $R_j^f(t) = n_j^f(t) r_j^f(t) = \frac{n_j^f(t)}{N(t)} c_r\left(g_j^f(t), \epsilon_j^f(t)\right)$. The third is the intensive-margin research labor to choose target quality $z^*$, $\frac{1}{N(t)} c_z(z^* \bar{x})$. Therefore, the market clearing condition for research labor at any time $t$ is

$$\begin{aligned}
\overline{R} &= \int_{(f,j)\in\mathcal{F}\times\mathcal{J}} \mathbb{1}_{\{\epsilon_j^f(t)\geq\epsilon^*(g,\bar{z},w_R)\}} \left[\frac{n_j^f(t)}{N(t)}\kappa + R_j^f(t) + g_j^f(t)\frac{n_j^f(t)}{N(t)} c_z(z^*\bar{x})\right] \\
&= \int_{(f,j)\in\mathcal{F}\times\mathcal{J}} s_j^f(t) \mathbb{1}_{\{\epsilon_j^f(t)\geq\epsilon^*(g,\bar{z},w_R)\}} \left[\kappa + c_r\left(g^*\left(\epsilon_j^f(t);\, g,\bar{z},w_R\right), \epsilon_j^f(t)\right) + g^*(\epsilon_j^f(t);\, g,\bar{z},w_R) c_z(\overline{zx})\right].
\end{aligned} \tag{B.3}$$

Inside the bracket of eq. (B.3), $\kappa$ is the lump sum innovation cost per knowledge share; expression $c_r\left(g^*\left(\epsilon_j^f(t);\, g,\bar{z},w_R\right), \epsilon_j^f(t)\right)$ is the optimal extensive-margin research labor hired to produce new knowledge at rate $g^*(\cdot)$; $c_z(\overline{zx})$ is the optimal intensive-margin research labor hired to get the optimal target quality $\bar{z}$ of each new patent in eq. (29). Let $\epsilon^*(g,\bar{z},w_R) = \epsilon^*$. The firm-sector states that are relevant for research labor aggregation are knowledge stock share $s_j^f(t)$ and innovation efficiency $\epsilon_j^f(t)$. Let $\Phi(s,\epsilon,t)$ denote the time-$t$ measure of firm-sector pairs with knowledge stock shares no higher than $s$ and innovation efficiency no higher than $\epsilon$. Aggregation at time $t$ by $(f,j)$ is then equivalent to aggregation by $(s,\epsilon)$ using $\Phi(s,\epsilon,t)$. By assumption, $\epsilon_j^f(t)$ is i.i.d., and so we can write $\Phi(s,\epsilon,t) = \Phi(s,t)G(\epsilon)$ with slight abuse of notation. By definition of knowledge stock shares, $\int_{(f,j)\in\mathcal{F}\times\mathcal{J}} s_j^f(t) = \int_s s\Phi(\mathrm{d}s,t) = 1$. Applying a law of large numbers, the RHS of eq. (B.3) becomes

$$\begin{aligned}
&RHS \\
&= \int_s s \int_{\epsilon\geq\epsilon^*} \left[\kappa + \frac{\gamma}{A}\left(\epsilon\mathcal{G}\left(\bar{z},w_R\right)\right)^{\frac{1}{\gamma}}(\epsilon)^{-\frac{1-\gamma}{\gamma}} + \epsilon\mathcal{G}(g,\bar{z},w_R)c_z(\overline{zx})\right] \mathrm{d}G(\epsilon)\,\Phi(\mathrm{d}s,t) \\
&= [1-G(\epsilon^*)]\kappa + g\cdot c_z(\overline{zx}) + \gamma\left(\int_{\epsilon^*}^{\infty}\epsilon\mathrm{d}G\right)\overline{A}^{\frac{\gamma}{1-\gamma}}\left(\frac{\mathcal{B}(g,\bar{z},w_R)}{w_R}\right)^{\frac{1}{1-\gamma}}.
\end{aligned}$$

Using eq. (26), the last term becomes $\frac{\gamma}{w_R(1-\gamma)} v\left(\int_{\epsilon^*}^{\infty}\epsilon\mathrm{d}G;\, g,\bar{z},w_R\right) = \frac{\gamma\mathcal{V}(g,\bar{z},w_R)}{w_R(1-\gamma)}\int_{\epsilon^*}^{\infty}\epsilon\mathrm{d}G$. Then $\overline{R} = RHS$ is eq. (30). Note that the explicit form and evolution rule of the distribution $\Phi(s,t)$ need not be part of the equilibrium definition.

## B.6 Details of subsection 4.6 and proof of Proposition 5

### B.6.1 The value of an existing patent

Consider an existing patent named $\omega \in \mathcal{N}(t)$ at time $t$. Its initial quality and age are fully summarized by the time-$t$ current quality, denoted as

$$z(\omega, t) = \widetilde{z} = z(\omega)e^{-\delta\tau(\omega,t)}.$$

Its current indegree is $d(\omega, t) = d$.

The time-$t$ present value of the associated production line is straightforward, given as

$$\int_0^\infty e^{-\rho\tau} \frac{\pi\widetilde{z}e^{-\delta\tau}}{N(t+\tau)\overline{z^o}} = \frac{\widetilde{z}}{N(t)}\mathcal{P}(g, \bar{z}).$$

The expected law of motion of $\omega$'s indegree in the future determines the citation payment flow, and it can be written as a differential equation identical to eq. (B.1), or

$$\frac{\partial k(\widetilde{z}, t+\tau \mid t)}{\partial \tau} = \widetilde{z}e^{-\delta\tau} \frac{\bar{x}(1-\bar{r})}{\bar{z}/(g+\delta)} + k(\widetilde{z}, t+\tau \mid t)g\bar{r}$$

with a different boundary condition $k(\widetilde{z}, t) = d$. This non-zero boundary condition results in an additional term in the differential equation's solution, such that, conditional on current indegree $d(\omega, t) = d$, the expected number of citations $\omega$ accumulates at $t + \tau$ is given by

$$k(\widetilde{z}, t+\tau \mid t) = k\left(\widetilde{z}, \tau\right) + de^{g\bar{r}\tau},$$

where $k(\cdot)$ is the expected-indegree function defined in eq. (15). The first-order derivative with respect to $\tau$ yields $\omega$'s citation growth,

$$\frac{\partial k(\widetilde{z}, t+\tau \mid t)}{\partial \tau} = \frac{\partial k(\widetilde{z}, \tau)}{\partial \tau} + dg\bar{r}e^{g\bar{r}\tau}.$$

The present value of future citation payment flows attracted to $\omega$ is then

$$\int_0^\infty e^{-\rho\tau} \frac{\phi}{N(t+\tau)} \left( \frac{\partial k(\widetilde{z}, \tau)}{\partial \tau} + dg\bar{r}e^{g\bar{r}\tau} \right) \mathrm{d}\tau = \frac{1}{N(t)} \left( \widetilde{z}\,\phi\bar{x}\mathcal{D}(g, \bar{z}) + d\phi \frac{g\bar{r}}{\rho + g(1-\bar{r})} \right).$$

The expression of $v(\widetilde{z}, d;\ g, \bar{z}, w_R)$ in eq. (31) follows immediately.

### B.6.2 Proof of Proposition 5

To ease expression, we drop the equilibrium index $(g, \bar{z}, w_R)$. Suppose $\mathcal{N}^f(t)$ is the set of all patents that firm $f$ already owns at $t$. By definition, $\bar{z}^f(t) = \int_{\omega\in\mathcal{N}^f(t)} z(\omega, t)/n^f(t)$, and $\bar{d}^f(t)$ is similar. The value of firm $f$'s existing pool of patents is

$$
\begin{aligned}
&\frac{1}{N(t)}\int_{\omega\in\mathcal{N}^f(t)} v(z(\omega,t),d(\omega,t);\ g,\bar{z},w_R)\\
=&\frac{1}{N(t)}\int_{\omega\in\mathcal{N}^f(t)}\left(z(\omega,t)\mathcal{P}+z(\omega,t)\phi\bar{x}\mathcal{D}+d(\omega,t)\phi\frac{g\bar{r}}{\rho+g(1-\bar{r})}\right)\\
=&\frac{1}{N(t)}\left(n^f(t)\bar{z}^f(t)\mathcal{P}+n^f(t)\bar{z}^f(t)\phi\bar{x}\mathcal{D}+n^f(t)\bar{d}^f(t)\phi\frac{g\bar{r}}{\rho+g(1-\bar{r})}\right)\\
=&s\left(\tilde{z}\mathcal{P}+\tilde{z}\phi\bar{x}\mathcal{D}+d\phi\frac{g\bar{r}}{\rho+g(1-\bar{r})}\right).
\end{aligned}
$$

The value of firm $f$ due to future innovation is as follows. For any future time $t+\tau$, the expected net flow value of firm $f$'s optimal innovation decision can be written as

$$
\begin{aligned}
&\mathbb{E}_t\left[\sum_{j\in\mathcal{J}} s_j^f(t+\tau)\mathbb{1}_{\{\epsilon_j^f(t+\tau)\geq\epsilon^*\}}\left(\epsilon_j^f(t+\tau)\mathcal{V}-w_R\kappa\right)\right]\\
=&\mathbb{E}_t\left[s^f(t+\tau)\right]\left(\mathcal{V}\int_{\epsilon^*}^{\infty}\epsilon dG-\left[1-G(\epsilon^*)\right]w_R\kappa\right)\\
=&\mathbb{E}_t\left[s^f(t+\tau)\right]\left(\mathcal{V}^*-\left[1-G(\epsilon^*)\right]w_R\kappa\right),
\end{aligned}
$$

where $\mathcal{V}(\cdot)$ is defined in eq. (26). Along a stationary equilibrium path, a firm's knowledge stock and the aggregate knowledge stock are expected to grow at the same rate $g$. Therefore, for any $\tau>0$, $\mathbb{E}_t\left[s^f(t+\tau)\right]=s^f(t)=s$. The present value of the payoff stream due to firm innovation becomes

$$
\int_0^{\infty}e^{-\rho\tau}s\Big(\mathcal{V}^*-\left[1-G(\epsilon^*)\right]w_R\kappa\Big)=\frac{s}{\rho}\Big(\mathcal{V}^*-\left[1-G(\epsilon^*)\right]w_R\kappa\Big).
$$

Proposition 5 follows.

# Appendix C Tables and Figures

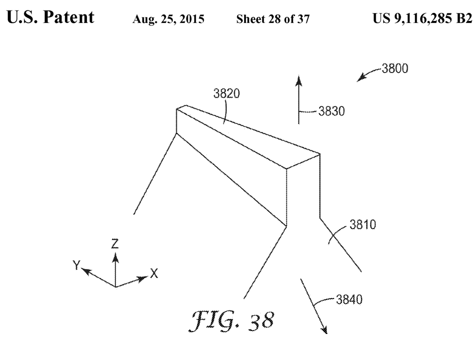


Figure 2: Diamond Cutting Tool

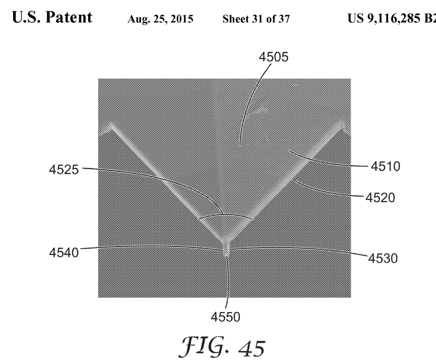


Figure 3: Light Directing Film

Table 2: Innovation Rate for All Firms

| | (1)<br>$\tau$ =‘ | (2)<br>$\tau$ =2 | (3)<br>$\tau$ =3 | (4)<br>$\tau$ =4 | (5)<br>$\tau$ =5 |
|---|---|---|---|---|---|
| lps | -1.005*** | -1.040*** | -1.075*** | -1.105*** | -1.130*** |
| | (-332.22) | (-298.92) | (-277.85) | (-261.01) | (-246.11) |
| lnonclass | 0.612*** | 0.617*** | 0.592*** | 0.558*** | 0.528*** |
| | (129.61) | (121.77) | (108.06) | (94.96) | (83.66) |
| lwpf_self | 0.723*** | 0.574*** | 0.476*** | 0.450*** | 0.450*** |
| | (11.95) | (9.99) | (8.19) | (7.76) | (7.70) |
| lwpf_peer | 0.0308*** | 0.0380*** | 0.0360*** | 0.0448*** | 0.0407*** |
| | (3.64) | (4.72) | (4.46) | (5.44) | (4.81) |
| lwpr_self | 0.993*** | 1.104*** | 1.131*** | 1.105*** | 1.112*** |
| | (10.15) | (10.90) | (10.50) | (9.85) | (9.43) |
| lwpr_peer | 0.0736*** | 0.0469*** | 0.0466*** | 0.0434*** | 0.0497*** |
| | (9.56) | (6.40) | (6.34) | (5.76) | (6.41) |
| $N$ | 483981 | 581050 | 599214 | 584269 | 551834 |
| $R^2$ | 0.819 | 0.791 | 0.789 | 0.795 | 0.805 |

Firm and year fixed effects included. Observations clustered by firm.

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 3: Extensive Innovation Rate for All Firms

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| lps | -0.372*** | -0.798*** | -1.283*** | -1.809*** | -2.331*** |
| | (-82.68) | (-76.28) | (-63.55) | (-58.55) | (-56.93) |
| lnonclass | -0.166*** | -0.355*** | -0.537*** | -0.710*** | -0.872*** |
| | (-52.32) | (-47.57) | (-36.98) | (-30.90) | (-28.59) |
| lwpf_self | 0.179*** | 0.369*** | 0.562*** | 0.882*** | 1.148** |
| | (4.72) | (4.56) | (3.82) | (3.46) | (3.27) |
| lwpf_peer | 0.252*** | 0.475*** | 0.596*** | 0.663*** | 0.718*** |
| | (42.00) | (36.09) | (26.59) | (18.86) | (14.20) |
| lwpr_self | 0.600*** | 1.299*** | 2.198*** | 3.254*** | 4.300*** |
| | (12.68) | (12.56) | (11.88) | (11.17) | (10.21) |
| lwpr_peer | 0.431*** | 0.797*** | 1.025*** | 1.169*** | 1.296*** |
| | (76.26) | (64.29) | (47.56) | (35.29) | (27.76) |
| $N$ | 576222 | 458316 | 394187 | 341563 | 295485 |
| $R^2$ | 0.443 | 0.512 | 0.534 | 0.530 | 0.538 |

Firm and year fixed effects included. Observations clustered by firm.

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 4: Intensive Innovation Rate for All Firms

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| lps | -0.544*** | -0.121*** | 0.356*** | 0.858*** | 1.362*** |
| | (-94.37) | (-11.12) | (17.15) | (27.42) | (33.29) |
| lnonclass | 0.583*** | 0.745*** | 0.871*** | 1.000*** | 1.124*** |
| | (103.65) | (97.82) | (64.26) | (46.62) | (39.46) |
| lwpf_self | 0.709*** | 0.430*** | 0.110 | -0.202 | -0.484 |
| | (10.79) | (5.01) | (0.78) | (-0.92) | (-1.60) |
| lwpf_peer | -0.164*** | -0.343*** | -0.430*** | -0.489*** | -0.547*** |
| | (-15.41) | (-21.39) | (-18.06) | (-13.94) | (-11.32) |
| lwpr_self | 0.345*** | -0.406*** | -1.428*** | -2.692*** | -3.831*** |
| | (3.91) | (-4.04) | (-9.11) | (-10.68) | (-10.05) |
| lwpr_peer | -0.319*** | -0.641*** | -0.844*** | -0.991*** | -1.119*** |
| | (-31.42) | (-42.62) | (-36.70) | (-29.99) | (-25.05) |
| $N$ | 342891 | 390133 | 366781 | 328697 | 288790 |
| $R^2$ | 0.684 | 0.469 | 0.389 | 0.381 | 0.403 |

Firm and year fixed effects included. Observations clustered by firm.

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 5: Growth in Market Value

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| lwpf_self | 0.0464*** | 0.0253* | 0.0299* | 0.0351* | 0.0348* | 0.0188* | 0.00404 | 0.00804 | 0.0160 | 0.00846 |
| | (5.71) | (2.33) | (2.42) | (2.44) | (2.11) | (2.42) | (0.46) | (0.87) | (1.75) | (0.91) |
| lwpf_peer | 0.0402** | 0.0119 | 0.0519 | 0.0807* | 0.0750 | 0.0331* | -0.00763 | 0.0197 | 0.0314 | -0.00590 |
| | (3.22) | (0.67) | (1.91) | (2.26) | (1.52) | (2.51) | (-0.46) | (1.01) | (1.38) | (-0.21) |
| lwpr_self | 0.0609*** | 0.0610*** | 0.0483*** | 0.0540*** | 0.0654*** | 0.00780 | 0.00922 | -0.00396 | -0.00403 | 0.00129 |
| | (6.61) | (5.39) | (3.54) | (3.55) | (3.54) | (1.12) | (1.08) | (-0.42) | (-0.52) | (0.12) |
| lwpr_peer | 0.0211 | 0.0572* | 0.0522 | 0.0370 | 0.0484 | -0.00249 | 0.000733 | -0.00658 | 0.0321 | -0.00178 |
| | (1.76) | (2.56) | (1.22) | (0.85) | (0.80) | (-0.19) | (0.04) | (-0.18) | (1.16) | (-0.05) |
| ltsm | 0.706*** | 0.573*** | 0.448*** | 0.418*** | 0.401*** | 0.793*** | 0.424*** | 0.164*** | 0.137*** | 0.0407 |
| | (29.65) | (19.40) | (12.84) | (10.20) | (7.99) | (31.30) | (14.13) | (5.17) | (4.62) | (1.28) |
| ltcw | -0.633*** | -0.492*** | -0.381*** | -0.350*** | -0.323*** | -0.783*** | -0.411*** | -0.169*** | -0.113*** | -0.0505 |
| | (-27.75) | (-17.93) | (-11.51) | (-9.07) | (-6.73) | (-31.61) | (-12.89) | (-4.99) | (-3.43) | (-1.46) |
| $N$ | 9146 | 7394 | 5928 | 4711 | 3668 | 8680 | 7090 | 5706 | 4535 | 3514 |
| $R^2$ | 0.376 | 0.334 | 0.332 | 0.366 | 0.392 | 0.627 | 0.690 | 0.752 | 0.815 | 0.859 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

lwpf is log scaled patent stock weighted number of new outward citations from firm's technology space.

lwpr is log scaled patent stock weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and market value are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Growth in Employment

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau=1$ | $\tau=2$ | $\tau=3$ | $\tau=4$ | $\tau=5$ | $\tau=1$ | $\tau=2$ | $\tau=3$ | $\tau=4$ | $\tau=5$ |
| lwpf_self | 0.0146*** | 0.0224*** | 0.0250*** | 0.0257*** | 0.0277*** | 0.00874*** | 0.0127*** | 0.0112*** | 0.00960*** | 0.00905** |
| | (10.44) | (10.61) | (9.17) | (7.67) | (7.15) | (6.37) | (6.49) | (4.97) | (3.55) | (3.03) |
| lwpf_peer | 0.0135*** | 0.0203*** | 0.0212*** | 0.0254*** | 0.0190** | 0.00995*** | 0.0117** | 0.00950* | 0.0101* | 0.00308 |
| | (5.37) | (5.29) | (4.28) | (4.31) | (2.98) | (4.17) | (3.17) | (2.11) | (1.99) | (0.54) |
| lwpr_self | 0.0134*** | 0.0190*** | 0.0210*** | 0.0243*** | 0.0259*** | 0.00986*** | 0.0137*** | 0.0148*** | 0.0173*** | 0.0181*** |
| | (10.07) | (9.10) | (7.67) | (7.47) | (6.86) | (7.78) | (7.12) | (6.15) | (6.27) | (5.68) |
| lwpr_peer | 0.0150*** | 0.0231*** | 0.0260*** | 0.0250*** | 0.0302*** | 0.0112*** | 0.0157*** | 0.0159*** | 0.0134** | 0.0160** |
| | (5.42) | (5.72) | (5.04) | (3.98) | (4.48) | (3.87) | (3.98) | (3.52) | (2.76) | (3.23) |
| ltsm | 0.0511*** | 0.0759*** | 0.0878*** | 0.0945*** | 0.0972*** | 0.0879*** | 0.121*** | 0.117*** | 0.107*** | 0.0948*** |
| | (12.66) | (11.38) | (10.28) | (9.24) | (8.28) | (11.93) | (10.42) | (8.46) | (6.81) | (5.50) |
| ltcw | -0.106*** | -0.145*** | -0.155*** | -0.159*** | -0.154*** | -0.140*** | -0.168*** | -0.152*** | -0.134*** | -0.108*** |
| | (-23.86) | (-20.25) | (-16.67) | (-13.95) | (-11.54) | (-21.16) | (-16.89) | (-12.60) | (-9.57) | (-6.83) |
| $N$ | 35419 | 32060 | 29043 | 26399 | 24022 | 34599 | 31437 | 28541 | 25987 | 23681 |
| $R^2$ | 0.126 | 0.158 | 0.180 | 0.199 | 0.218 | 0.339 | 0.446 | 0.523 | 0.583 | 0.627 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

lwpf is log scaled patent stock weighted number of new outward citations from firm's technology space.

lwpr is log scaled patent stock weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen and capital are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Growth in Capital

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| lwpf_self | 0.0160*** | 0.0239*** | 0.0285*** | 0.0315*** | 0.0338*** | 0.00826*** | 0.0119*** | 0.0113*** | 0.0110*** | 0.0103** |
| | (11.00) | (10.46) | (9.40) | (8.49) | (7.74) | (6.22) | (5.98) | (4.65) | (3.90) | (3.20) |
| lwpf_peer | 0.0152*** | 0.0241*** | 0.0255*** | 0.0318*** | 0.0259*** | 0.00773** | 0.00924* | 0.00801 | 0.0117* | 0.00907 |
| | (6.13) | (5.70) | (4.91) | (5.17) | (3.86) | (3.27) | (2.46) | (1.74) | (2.24) | (1.58) |
| lwpr_self | 0.0152*** | 0.0237*** | 0.0275*** | 0.0307*** | 0.0347*** | 0.00999*** | 0.0148*** | 0.0168*** | 0.0173*** | 0.0191*** |
| | (9.97) | (9.68) | (8.62) | (8.01) | (7.92) | (8.12) | (8.05) | (7.26) | (6.18) | (6.04) |
| lwpr_peer | 0.0167*** | 0.0260*** | 0.0350*** | 0.0342*** | 0.0446*** | 0.00780** | 0.0104** | 0.0125** | 0.00977 | 0.0139* |
| | (6.02) | (6.15) | (6.31) | (4.97) | (5.91) | (2.84) | (2.78) | (2.75) | (1.89) | (2.56) |
| ltsm | 0.0545*** | 0.0933*** | 0.114*** | 0.122*** | 0.126*** | 0.0800*** | 0.145*** | 0.164*** | 0.160*** | 0.152*** |
| | (11.29) | (10.71) | (9.95) | (9.07) | (8.25) | (10.73) | (10.82) | (9.66) | (8.38) | (7.41) |
| ltcw | -0.112*** | -0.175*** | -0.206*** | -0.218*** | -0.221*** | -0.134*** | -0.198*** | -0.210*** | -0.198*** | -0.179*** |
| | (-22.24) | (-20.10) | (-17.92) | (-15.64) | (-13.62) | (-20.36) | (-18.24) | (-15.18) | (-12.25) | (-9.95) |
| $N$ | 35669 | 32310 | 29263 | 26591 | 24175 | 34842 | 31686 | 28764 | 26187 | 23827 |
| $R^2$ | 0.165 | 0.223 | 0.255 | 0.276 | 0.289 | 0.401 | 0.521 | 0.599 | 0.652 | 0.692 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

lwpf is log scaled patent stock weighted number of new outward citations from firm's technology space.

lwpr is log scaled patent stock weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen and capital are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Growth in Profit

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau=1$ | $\tau=2$ | $\tau=3$ | $\tau=4$ | $\tau=5$ | $\tau=1$ | $\tau=2$ | $\tau=3$ | $\tau=4$ | $\tau=5$ |
| lwpf_self | 0.0156$^{***}$ | 0.0259$^{***}$ | 0.0314$^{***}$ | 0.0336$^{***}$ | 0.0352$^{***}$ | 0.00851$^{***}$ | 0.0139$^{***}$ | 0.0142$^{***}$ | 0.0146$^{***}$ | 0.0119$^{**}$ |
| | (7.65) | (8.88) | (8.92) | (8.18) | (7.67) | (4.37) | (5.33) | (4.84) | (4.51) | (3.29) |
| lwpf_peer | 0.0128$^{**}$ | 0.0166$^{**}$ | 0.0101 | 0.0209$^{**}$ | 0.0109 | 0.00722 | 0.00227 | -0.00470 | 0.00606 | -0.00244 |
| | (3.00) | (2.71) | (1.55) | (2.93) | (1.52) | (1.78) | (0.40) | (-0.81) | (0.97) | (-0.39) |
| lwpr_self | 0.0199$^{***}$ | 0.0266$^{***}$ | 0.0253$^{***}$ | 0.0296$^{***}$ | 0.0334$^{***}$ | 0.0159$^{***}$ | 0.0194$^{***}$ | 0.0162$^{***}$ | 0.0179$^{***}$ | 0.0213$^{***}$ |
| | (10.17) | (9.49) | (7.25) | (7.55) | (7.36) | (9.08) | (7.83) | (5.55) | (5.55) | (5.70) |
| lwpr_peer | 0.0180$^{***}$ | 0.0261$^{***}$ | 0.0350$^{***}$ | 0.0288$^{***}$ | 0.0367$^{***}$ | 0.00958$^{*}$ | 0.0115 | 0.0157$^{**}$ | 0.0103 | 0.0161$^{**}$ |
| | (4.10) | (4.15) | (5.25) | (3.91) | (4.64) | (2.13) | (1.89) | (2.80) | (1.77) | (2.73) |
| ltsm | 0.0883$^{***}$ | 0.114$^{***}$ | 0.115$^{***}$ | 0.114$^{***}$ | 0.110$^{***}$ | 0.164$^{***}$ | 0.189$^{***}$ | 0.154$^{***}$ | 0.130$^{***}$ | 0.117$^{***}$ |
| | (12.32) | (10.84) | (9.59) | (8.45) | (7.52) | (12.80) | (10.93) | (8.97) | (7.45) | (6.54) |
| ltcw | -0.133$^{***}$ | -0.166$^{***}$ | -0.161$^{***}$ | -0.155$^{***}$ | -0.144$^{***}$ | -0.207$^{***}$ | -0.229$^{***}$ | -0.184$^{***}$ | -0.154$^{***}$ | -0.134$^{***}$ |
| | (-18.57) | (-15.43) | (-12.86) | (-10.68) | (-8.92) | (-19.31) | (-15.38) | (-11.95) | (-9.37) | (-7.65) |
| $N$ | 33475 | 30445 | 27707 | 25284 | 23099 | 32699 | 29860 | 27231 | 24909 | 22781 |
| $R^2$ | 0.150 | 0.184 | 0.203 | 0.232 | 0.249 | 0.387 | 0.472 | 0.536 | 0.588 | 0.633 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

lwpf is log scaled patent stock weighted number of new outward citations from firm's technology space.

lwpr is log scaled patent stock weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and profit are included in the regression, but not reported.

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 9: Growth in Sales

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| lwpf_self | 0.0163*** | 0.0252*** | 0.0297*** | 0.0304*** | 0.0311*** | 0.0102*** | 0.0150*** | 0.0146*** | 0.0128*** | 0.0105** |
| | (8.60) | (9.18) | (8.57) | (7.62) | (6.84) | (6.14) | (6.30) | (5.23) | (4.31) | (3.24) |
| lwpf_peer | 0.0160*** | 0.0232*** | 0.0239*** | 0.0248*** | 0.0177** | 0.00510 | 0.00415 | 0.00497 | 0.00428 | -0.000551 |
| | (5.02) | (4.79) | (4.16) | (3.84) | (2.68) | (1.65) | (1.01) | (1.02) | (0.81) | (-0.10) |
| lwpr_self | 0.0190*** | 0.0276*** | 0.0290*** | 0.0327*** | 0.0348*** | 0.0154*** | 0.0213*** | 0.0207*** | 0.0218*** | 0.0230*** |
| | (10.52) | (10.17) | (8.59) | (8.37) | (7.79) | (9.49) | (9.22) | (7.58) | (6.97) | (6.44) |
| lwpr_peer | 0.0151*** | 0.0163** | 0.0200** | 0.0207** | 0.0265*** | 0.00900** | 0.00463 | 0.00509 | 0.00780 | 0.00886 |
| | (4.70) | (3.15) | (3.13) | (2.87) | (3.60) | (2.88) | (1.03) | (1.00) | (1.50) | (1.72) |
| ltsm | 0.0673*** | 0.0946*** | 0.105*** | 0.106*** | 0.104*** | 0.125*** | 0.163*** | 0.150*** | 0.122*** | 0.106*** |
| | (12.05) | (10.76) | (9.78) | (8.94) | (8.09) | (12.07) | (10.39) | (8.82) | (7.45) | (6.48) |
| ltcw | -0.134*** | -0.181*** | -0.194*** | -0.192*** | -0.179*** | -0.187*** | -0.223*** | -0.202*** | -0.169*** | -0.132*** |
| | (-20.59) | (-18.12) | (-15.79) | (-13.82) | (-11.74) | (-19.75) | (-15.94) | (-13.19) | (-10.76) | (-8.09) |
| $N$ | 35424 | 32123 | 29130 | 26483 | 24094 | 34608 | 31506 | 28635 | 26084 | 23754 |
| $R^2$ | 0.165 | 0.202 | 0.226 | 0.242 | 0.251 | 0.411 | 0.526 | 0.592 | 0.644 | 0.677 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

lwpf is log scaled patent stock weighted number of new outward citations from firm’s technology space.

lwpr is log scaled patent stock weighted number of new inward citations to firm’s technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and sale are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Growth in TFP

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| lwpf_self | 0.00258 | 0.00343 | 0.00669** | 0.00905*** | 0.00685* | 0.000591 | 0.000174 | 0.00348 | 0.00414 | 0.00231 |
| | (1.49) | (1.61) | (2.89) | (3.76) | (2.54) | (0.34) | (0.08) | (1.63) | (1.93) | (0.96) |
| lwpf_peer | -0.00447 | -0.00671 | -0.00459 | -0.00625 | -0.00521 | -0.00540 | -0.00715* | -0.00331 | -0.00527 | -0.00254 |
| | (-1.18) | (-1.62) | (-0.82) | (-1.65) | (-1.29) | (-1.57) | (-1.97) | (-0.61) | (-1.42) | (-0.77) |
| lwpr_self | 0.00446** | 0.00296 | 0.00463* | 0.00654** | 0.00735** | 0.00432** | 0.00277 | 0.00468* | 0.00759*** | 0.00781*** |
| | (3.12) | (1.48) | (2.13) | (2.60) | (2.98) | (3.09) | (1.47) | (2.40) | (3.63) | (3.83) |
| lwpr_peer | 0.00689 | 0.00587 | 0.00625 | 0.00454 | -0.00183 | -0.000369 | 0.000795 | -0.000248 | -0.00301 | -0.0103** |
| | (1.65) | (1.36) | (1.29) | (1.02) | (-0.39) | (-0.11) | (0.20) | (-0.06) | (-0.80) | (-2.62) |
| ltsm | 0.0332*** | 0.0417*** | 0.0400*** | 0.0394*** | 0.0391*** | 0.0454*** | 0.0430*** | 0.0345*** | 0.0297*** | 0.0260** |
| | (9.57) | (8.84) | (7.49) | (6.62) | (6.21) | (8.03) | (6.26) | (4.84) | (3.77) | (3.00) |
| ltcw | -0.0254*** | -0.0251*** | -0.0202** | -0.0162* | -0.0119 | -0.0441*** | -0.0378*** | -0.0274*** | -0.0272** | -0.0235* |
| | (-6.10) | (-4.48) | (-3.09) | (-2.21) | (-1.50) | (-7.42) | (-5.30) | (-3.47) | (-3.16) | (-2.52) |
| $N$ | 25306 | 22912 | 20917 | 19128 | 17529 | 24842 | 22534 | 20597 | 18864 | 17278 |
| $R^2$ | 0.172 | 0.244 | 0.280 | 0.302 | 0.317 | 0.348 | 0.467 | 0.522 | 0.560 | 0.595 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

lwpf is log scaled patent stock weighted number of new outward citations from firm's technology space.

lwpr is log scaled patent stock weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and tfp are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Citation Count across Categories

| | (1) lncite | (2) lncite | (3) lncite | (4) lncite |
|---|---|---|---|---|
| lcps | 0.277*** | 0.653*** | 0.277*** | 0.653*** |
| | (36.15) | (92.67) | (36.15) | (92.67) |
| ldps | 0.302*** | 0.579*** | 0.302*** | 0.579*** |
| | (34.62) | (72.60) | (34.62) | (72.60) |
| _cons | -3.701*** | -9.534*** | | |
| | (-36.42) | (-118.63) | | |
| $N$ | 1531713 | 1512213 | 1531713 | 1512213 |
| $R^2$ | 0.148 | 0.758 | 0.148 | 0.758 |
| Citing-nclass FE | Yes | No | Yes | No |
| Cited-nclass FE | Yes | No | Yes | No |
| Year FE | Yes | Yes | Yes | Yes |
| Citing-nclass-Cited-nclass FE | No | Yes | No | Yes |

*t* statistics in parentheses

lncite is log scaled citation count from citing category to cited category.

lcps is log scaled patent stock in the citing category.

ldps is scaled patent stock in the cited category.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Quality Weighted Innovation Rate for All Firms

| | (1) $\tau=1$ | (2) $\tau=2$ | (3) $\tau=3$ | (4) $\tau=4$ | (5) $\tau=5$ |
|---|---|---|---|---|---|
| lcwps | -1.396*** | -3.296*** | -5.903*** | -9.453*** | -14.12*** |
| | (-44.52) | (-34.89) | (-27.77) | (-20.27) | (-16.13) |
| lnonclass | 0.397*** | 0.491*** | 0.202 | -0.694* | -2.280*** |
| | (21.51) | (10.31) | (1.47) | (-1.98) | (-3.29) |
| lwpf_self | -0.450** | -0.968** | -1.346 | -2.919 | -3.729 |
| | (-2.75) | (-2.82) | (-1.60) | (-1.24) | (-0.67) |
| lwpf_peer | 0.00125 | 0.132 | 0.153 | 0.258 | 0.375 |
| | (0.04) | (1.56) | (0.87) | (0.79) | (0.71) |
| lwpr_self | 2.244*** | 5.067*** | 7.739*** | 10.17*** | 10.88 |
| | (13.98) | (10.77) | (6.54) | (3.75) | (1.87) |
| lwpr_peer | 0.0276 | 0.200** | 0.612*** | 1.145** | 1.924** |
| | (1.00) | (2.80) | (3.63) | (3.02) | (2.83) |
| $N$ | 956556 | 849343 | 755730 | 672557 | 599414 |
| $R^2$ | 0.228 | 0.245 | 0.266 | 0.246 | 0.227 |

Firm and year fixed effects included. Observations clustered by firm.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Growth in Market Value

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| ldwpf_self | 0.0446*** | 0.0241* | 0.0285* | 0.0341* | 0.0337* | 0.0181* | 0.00382 | 0.00711 | 0.0150 | 0.00759 |
| | (5.83) | (2.36) | (2.46) | (2.51) | (2.18) | (2.48) | (0.47) | (0.82) | (1.75) | (0.86) |
| ldwpf_peer | 0.0344** | 0.0114 | 0.0482* | 0.0608 | 0.0425 | 0.0256* | -0.00823 | 0.0228 | 0.0192 | -0.0238 |
| | (3.00) | (0.70) | (2.02) | (1.83) | (0.92) | (2.12) | (-0.52) | (1.33) | (0.86) | (-0.91) |
| ldwpr_self | 0.114*** | 0.115*** | 0.0914*** | 0.102*** | 0.122*** | 0.0152 | 0.0180 | -0.00710 | -0.00642 | 0.00203 |
| | (6.77) | (5.50) | (3.61) | (3.62) | (3.53) | (1.18) | (1.14) | (-0.41) | (-0.46) | (0.11) |
| ldwpr_peer | 0.0251 | 0.0825 | 0.0718 | 0.0511 | 0.0586 | -0.0153 | -0.0157 | -0.0222 | 0.0781 | -0.0328 |
| | (1.06) | (1.82) | (0.82) | (0.60) | (0.48) | (-0.60) | (-0.39) | (-0.28) | (1.34) | (-0.45) |
| ltsm | 0.707*** | 0.574*** | 0.451*** | 0.422*** | 0.406*** | 0.794*** | 0.424*** | 0.164*** | 0.136*** | 0.0409 |
| | (29.65) | (19.42) | (12.88) | (10.24) | (8.05) | (31.28) | (14.12) | (5.17) | (4.56) | (1.28) |
| ltcw | -0.632*** | -0.491*** | -0.381*** | -0.352*** | -0.323*** | -0.782*** | -0.409*** | -0.169*** | -0.112*** | -0.0507 |
| | (-27.69) | (-17.86) | (-11.46) | (-9.08) | (-6.72) | (-31.56) | (-12.84) | (-4.99) | (-3.41) | (-1.46) |
| $N$ | 9146 | 7394 | 5928 | 4711 | 3668 | 8680 | 7090 | 5706 | 4535 | 3514 |
| $R^2$ | 0.376 | 0.333 | 0.332 | 0.365 | 0.391 | 0.627 | 0.690 | 0.752 | 0.815 | 0.859 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

ldwpf is log scaled patent stock and edge betweenness centrality double weighted number of new outward citations from firm's technology space.

ldwpr is log scaled patent stock and edge betweenness centrality double weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and market value are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: Growth in Employment

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| ldwpf_self | 0.0123*** | 0.0184*** | 0.0203*** | 0.0190*** | 0.0190*** | 0.00723*** | 0.0107*** | 0.00953*** | 0.00650* | 0.00458 |
| | (8.82) | (8.81) | (7.37) | (5.52) | (4.72) | (5.28) | (5.46) | (4.10) | (2.31) | (1.49) |
| ldwpf_peer | 0.0233*** | 0.0463*** | 0.0410*** | 0.0416** | 0.0303 | 0.0181*** | 0.0321*** | 0.0167 | 0.00651 | -0.00672 |
| | (5.72) | (5.94) | (3.83) | (2.81) | (1.57) | (4.49) | (4.42) | (1.71) | (0.58) | (-0.51) |
| ldwpr_self | 0.0186*** | 0.0240*** | 0.0249*** | 0.0278*** | 0.0291*** | 0.0109*** | 0.0132*** | 0.0140** | 0.0153** | 0.0168* |
| | (7.34) | (5.95) | (4.61) | (4.20) | (3.73) | (4.41) | (3.39) | (2.82) | (2.60) | (2.43) |
| ldwpr_peer | 0.0534*** | 0.0925*** | 0.143*** | 0.166*** | 0.167*** | 0.0458*** | 0.0649*** | 0.0907*** | 0.114*** | 0.0972** |
| | (5.58) | (5.54) | (5.97) | (4.88) | (3.80) | (4.34) | (3.85) | (3.68) | (3.97) | (2.91) |
| ltsm | 0.0528*** | 0.0782*** | 0.0904*** | 0.0975*** | 0.101*** | 0.0877*** | 0.120*** | 0.116*** | 0.107*** | 0.0947*** |
| | (13.23) | (11.94) | (10.76) | (9.66) | (8.67) | (11.93) | (10.40) | (8.43) | (6.76) | (5.47) |
| ltcw | -0.104*** | -0.142*** | -0.152*** | -0.155*** | -0.148*** | -0.140*** | -0.168*** | -0.152*** | -0.134*** | -0.106*** |
| | (-23.57) | (-20.18) | (-16.53) | (-13.64) | (-11.08) | (-21.10) | (-16.93) | (-12.55) | (-9.47) | (-6.72) |
| $N$ | 35424 | 32065 | 29048 | 26402 | 24025 | 34603 | 31442 | 28546 | 25990 | 23684 |
| $R^2$ | 0.125 | 0.158 | 0.180 | 0.199 | 0.216 | 0.340 | 0.446 | 0.522 | 0.582 | 0.626 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

ldwpf is log scaled patent stock and edge betweenness centrality double weighted number of new outward citations from firm's technology space.

ldwpr is log scaled patent stock and edge betweenness centrality double weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen and capital are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Growth in Capital

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| ldwpf_self | 0.0131*** | 0.0187*** | 0.0215*** | 0.0231*** | 0.0238*** | 0.00774*** | 0.0113*** | 0.0101*** | 0.00931** | 0.00833* |
| | (9.04) | (8.14) | (6.92) | (6.04) | (5.27) | (5.82) | (5.64) | (4.03) | (3.19) | (2.47) |
| ldwpf_peer | 0.0302*** | 0.0675*** | 0.0661*** | 0.0881*** | 0.0825*** | 0.0189*** | 0.0364*** | 0.0212 | 0.0229 | 0.0162 |
| | (7.46) | (7.74) | (4.88) | (4.85) | (3.58) | (4.82) | (4.93) | (1.84) | (1.72) | (0.98) |
| ldwpr_self | 0.0218*** | 0.0319*** | 0.0354*** | 0.0376*** | 0.0433*** | 0.0128*** | 0.0173*** | 0.0190*** | 0.0162* | 0.0190* |
| | (7.31) | (6.59) | (5.50) | (4.78) | (4.76) | (4.99) | (4.32) | (3.65) | (2.52) | (2.57) |
| ldwpr_peer | 0.0645*** | 0.118*** | 0.211*** | 0.224*** | 0.274*** | 0.0457*** | 0.0625*** | 0.109*** | 0.122*** | 0.125*** |
| | (6.20) | (6.43) | (7.10) | (5.28) | (5.25) | (4.31) | (3.95) | (4.18) | (4.00) | (3.45) |
| ltsm | 0.0564*** | 0.0957*** | 0.116*** | 0.125*** | 0.130*** | 0.0794*** | 0.143*** | 0.162*** | 0.159*** | 0.151*** |
| | (11.81) | (11.19) | (10.40) | (9.49) | (8.66) | (10.68) | (10.76) | (9.60) | (8.31) | (7.34) |
| ltcw | -0.111*** | -0.173*** | -0.204*** | -0.216*** | -0.218*** | -0.135*** | -0.199*** | -0.210*** | -0.199*** | -0.179*** |
| | (-22.14) | (-20.26) | (-18.04) | (-15.64) | (-13.57) | (-20.48) | (-18.41) | (-15.25) | (-12.28) | (-9.99) |
| $N$ | 35671 | 32313 | 29266 | 26593 | 24177 | 34844 | 31689 | 28767 | 26189 | 23829 |
| $R^2$ | 0.166 | 0.227 | 0.258 | 0.278 | 0.291 | 0.403 | 0.522 | 0.600 | 0.653 | 0.692 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

ldwpf is log scaled patent stock and edge betweenness centrality double weighted number of new outward citations from firm's technology space.

ldwpr is log scaled patent stock and edge betweenness centrality double weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen and capital are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 16: Growth in Profit

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| ldwpf_self | 0.0139*** | 0.0236*** | 0.0271*** | 0.0273*** | 0.0287*** | 0.00846*** | 0.0151*** | 0.0136*** | 0.0121*** | 0.0102** |
| | (6.78) | (7.91) | (7.55) | (6.35) | (5.93) | (4.28) | (5.69) | (4.54) | (3.65) | (2.79) |
| ldwpf_peer | 0.0288*** | 0.0517*** | 0.0328* | 0.0349 | 0.0125 | 0.0251*** | 0.0261** | 0.00684 | 0.00607 | -0.00613 |
| | (4.24) | (4.68) | (1.97) | (1.63) | (0.54) | (4.13) | (2.62) | (0.47) | (0.37) | (-0.35) |
| ldwpr_self | 0.0309*** | 0.0397*** | 0.0349*** | 0.0407*** | 0.0440*** | 0.0249*** | 0.0294*** | 0.0220*** | 0.0227** | 0.0268*** |
| | (8.22) | (7.08) | (4.91) | (5.04) | (4.70) | (7.33) | (5.90) | (3.57) | (3.27) | (3.40) |
| ldwpr_peer | 0.0745*** | 0.114*** | 0.216*** | 0.246*** | 0.293*** | 0.0535*** | 0.0647** | 0.115*** | 0.144*** | 0.163*** |
| | (5.53) | (4.71) | (6.09) | (5.24) | (5.39) | (4.44) | (3.12) | (3.72) | (3.91) | (3.74) |
| ltsm | 0.0904*** | 0.117*** | 0.118*** | 0.118*** | 0.115*** | 0.163*** | 0.188*** | 0.153*** | 0.130*** | 0.117*** |
| | (12.74) | (11.29) | (10.01) | (8.86) | (7.93) | (12.81) | (10.92) | (8.94) | (7.41) | (6.52) |
| ltcw | -0.133*** | -0.166*** | -0.161*** | -0.154*** | -0.142*** | -0.208*** | -0.231*** | -0.186*** | -0.155*** | -0.136*** |
| | (-18.70) | (-15.63) | (-13.06) | (-10.75) | (-8.86) | (-19.49) | (-15.56) | (-12.09) | (-9.44) | (-7.75) |
| $N$ | 33476 | 30447 | 27710 | 25287 | 23102 | 32700 | 29862 | 27233 | 24911 | 22783 |
| $R^2$ | 0.150 | 0.186 | 0.205 | 0.233 | 0.249 | 0.388 | 0.474 | 0.537 | 0.589 | 0.633 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

ldwpf is log scaled patent stock and edge betweenness centrality double weighted number of new outward citations from firm's technology space.

ldwpr is log scaled patent stock and edge betweenness centrality double weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and profit are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 17: Growth in Sales

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 | $\tau$ =1 | $\tau$ =2 | $\tau$ =3 | $\tau$ =4 | $\tau$ =5 |
| ldwpf_self | 0.0143*** | 0.0225*** | 0.0257*** | 0.0257*** | 0.0242*** | 0.0107*** | 0.0168*** | 0.0156*** | 0.0132*** | 0.00952** |
| | (7.49) | (8.02) | (7.11) | (6.22) | (5.12) | (6.40) | (6.94) | (5.42) | (4.37) | (2.95) |
| ldwpf_peer | 0.0329*** | 0.0601*** | 0.0649*** | 0.0652** | 0.0433 | 0.0185** | 0.0271** | 0.0304* | 0.0126 | -0.00683 |
| | (5.73) | (5.58) | (4.26) | (3.17) | (1.86) | (3.28) | (3.15) | (2.46) | (0.89) | (-0.43) |
| ldwpr_self | 0.0321*** | 0.0436*** | 0.0438*** | 0.0473*** | 0.0484*** | 0.0292*** | 0.0386*** | 0.0359*** | 0.0338*** | 0.0342*** |
| | (9.22) | (8.27) | (6.60) | (6.08) | (5.40) | (8.94) | (8.13) | (6.36) | (5.16) | (4.63) |
| ldwpr_peer | 0.0523*** | 0.0640** | 0.109** | 0.125** | 0.183*** | 0.0440*** | 0.0279 | 0.0430 | 0.0914** | 0.118** |
| | (4.26) | (2.60) | (3.01) | (2.60) | (3.31) | (3.48) | (1.22) | (1.34) | (2.60) | (3.02) |
| ltsm | 0.0692*** | 0.0972*** | 0.108*** | 0.110*** | 0.109*** | 0.125*** | 0.163*** | 0.149*** | 0.123*** | 0.106*** |
| | (12.56) | (11.24) | (10.23) | (9.37) | (8.52) | (12.10) | (10.39) | (8.81) | (7.44) | (6.47) |
| ltcw | -0.133*** | -0.180*** | -0.192*** | -0.190*** | -0.176*** | -0.188*** | -0.225*** | -0.204*** | -0.171*** | -0.133*** |
| | (-20.62) | (-18.21) | (-15.83) | (-13.76) | (-11.60) | (-19.99) | (-16.14) | (-13.34) | (-10.88) | (-8.19) |
| $N$ | 35428 | 32128 | 29135 | 26487 | 24098 | 34613 | 31510 | 28639 | 26087 | 23757 |
| $R^2$ | 0.165 | 0.203 | 0.228 | 0.242 | 0.251 | 0.413 | 0.527 | 0.593 | 0.644 | 0.677 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

ldwpf is log scaled patent stock and edge betweenness centrality double weighted number of new outward citations from firm's technology space.

ldwpr is log scaled patent stock and edge betweenness centrality double weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and sale are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 18: Growth in TFP

| | (1) $\tau$ =1 | (2) $\tau$ =2 | (3) $\tau$ =3 | (4) $\tau$ =4 | (5) $\tau$ =5 | (6) $\tau$ =1 | (7) $\tau$ =2 | (8) $\tau$ =3 | (9) $\tau$ =4 | (10) $\tau$ =5 |
|---|---|---|---|---|---|---|---|---|---|---|
| ldwpf_self | 0.00451* | 0.00473* | 0.00572* | 0.00672** | 0.00396 | 0.00397* | 0.00295 | 0.00450* | 0.00412 | 0.00140 |
| | (2.53) | (2.13) | (2.33) | (2.60) | (1.35) | (2.28) | (1.35) | (2.01) | (1.85) | (0.54) |
| ldwpf_peer | -0.00452 | 0.00554 | 0.00898 | -0.00196 | -0.0113 | -0.00620 | 0.0106 | 0.0218 | 0.0136 | 0.0156 |
| | (-0.50) | (0.46) | (0.52) | (-0.15) | (-0.75) | (-0.72) | (1.09) | (1.52) | (1.10) | (1.31) |
| ldwpr_self | 0.0128*** | 0.0107** | 0.0113** | 0.0123* | 0.0143** | 0.0153*** | 0.0135*** | 0.0140*** | 0.0172*** | 0.0177*** |
| | (4.50) | (2.69) | (2.65) | (2.43) | (2.82) | (5.40) | (3.54) | (3.57) | (3.96) | (4.08) |
| ldwpr_peer | 0.0463* | 0.0302 | 0.0609 | 0.0688* | 0.0746* | 0.0187 | 0.00184 | 0.0111 | 0.00412 | 0.00339 |
| | (2.24) | (1.16) | (1.87) | (2.28) | (2.17) | (1.10) | (0.08) | (0.42) | (0.15) | (0.11) |
| ltsm | 0.0338*** | 0.0420*** | 0.0407*** | 0.0404*** | 0.0396*** | 0.0460*** | 0.0429*** | 0.0344*** | 0.0299*** | 0.0260** |
| | (9.71) | (8.95) | (7.65) | (6.80) | (6.30) | (8.12) | (6.26) | (4.81) | (3.78) | (3.00) |
| ltcw | -0.0286*** | -0.0288*** | -0.0237*** | -0.0185* | -0.0141 | -0.0470*** | -0.0408*** | -0.0306*** | -0.0297*** | -0.0266** |
| | (-6.80) | (-5.11) | (-3.61) | (-2.52) | (-1.75) | (-7.80) | (-5.64) | (-3.84) | (-3.43) | (-2.81) |
| $N$ | 25308 | 22914 | 20918 | 19129 | 17529 | 24846 | 22534 | 20597 | 18864 | 17278 |
| $R^2$ | 0.176 | 0.247 | 0.284 | 0.305 | 0.318 | 0.353 | 0.468 | 0.523 | 0.560 | 0.595 |
| Industry FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Firm FE | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Observations are clustered at firm level.

ldwpf is log scaled patent stock and edge betweenness centrality double weighted number of new outward citations from firm's technology space.

ldwpr is log scaled patent stock and edge betweenness centrality double weighted number of new inward citations to firm's technology space.

self means new citation links made by self firm.

peer means new citation links made by peer firms.

log scaled employmen, capital and tfp are included in the regression, but not reported.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 19: Growth Volatility and Innovation Concentration across Sectors

| | (1) |
|---|---|
| | $\tau$ =5 |
| lps | -0.410*** |
| | (-9.89) |
| lnonclass | -0.0524 |
| | (-1.95) |
| herfindahl | 0.0676*** |
| | (3.61) |
| $N$ | 151813 |
| $R^2$ | 0.163 |

Firm and year fixed effects included. Observations clustered by firm.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

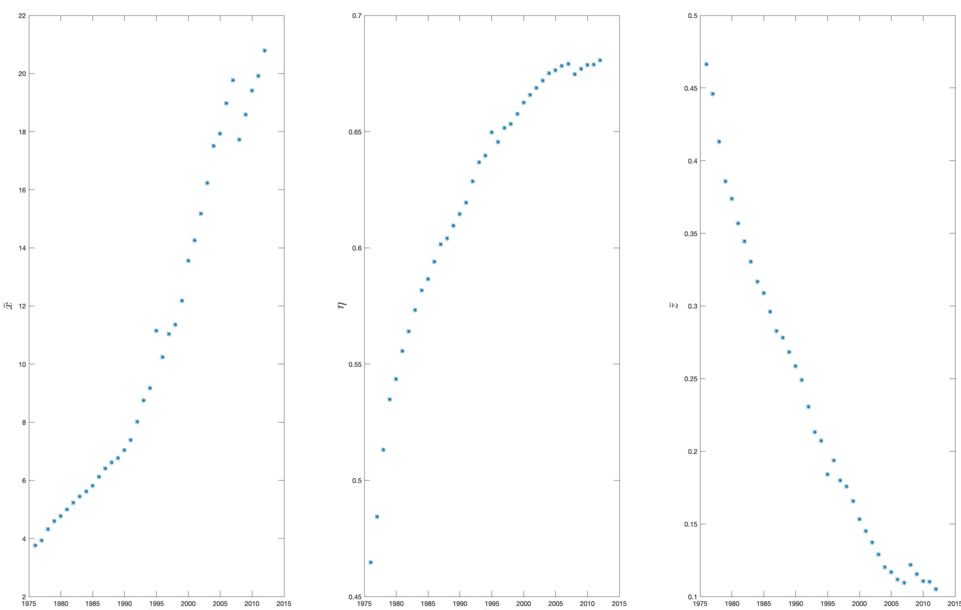


Figure 4: Patent Network Parameters

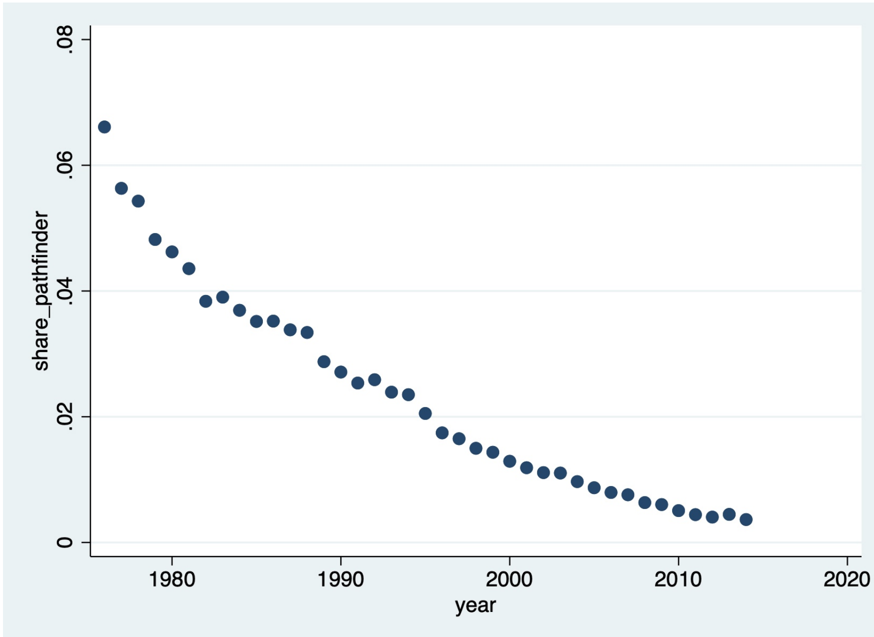


Figure 5: Share of Pathfinder Patents Over Time

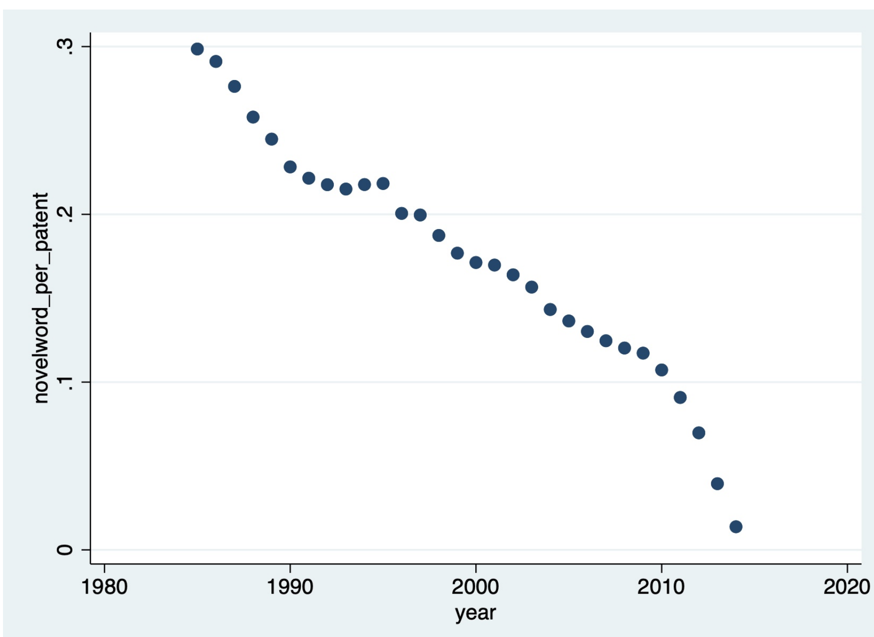


Figure 6: Novel Word per Patent Over Time

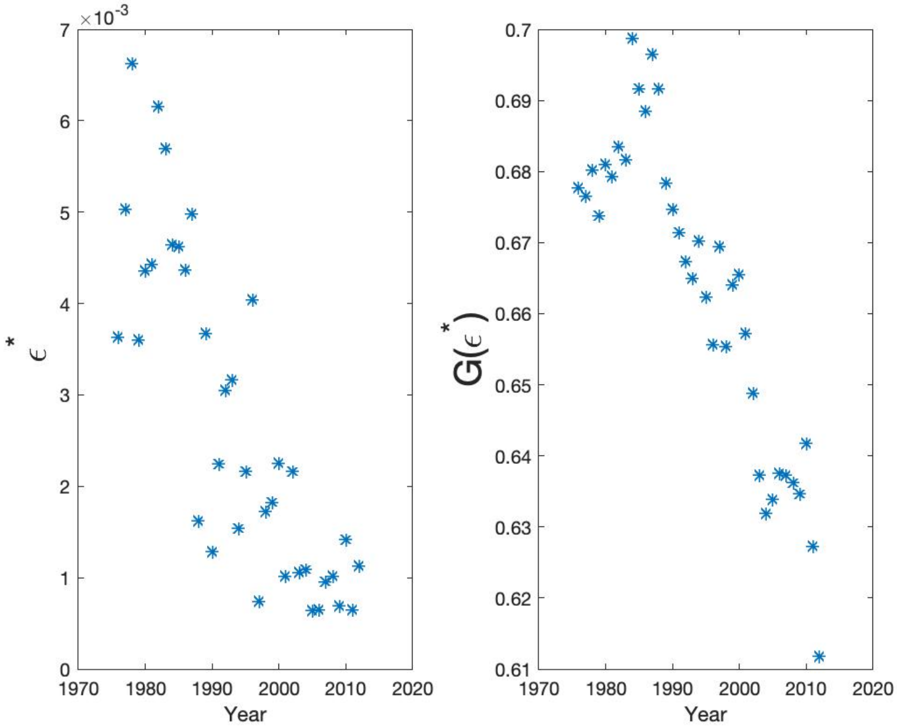


Figure 7: $G(\epsilon^*)_t$ and $\epsilon^*$

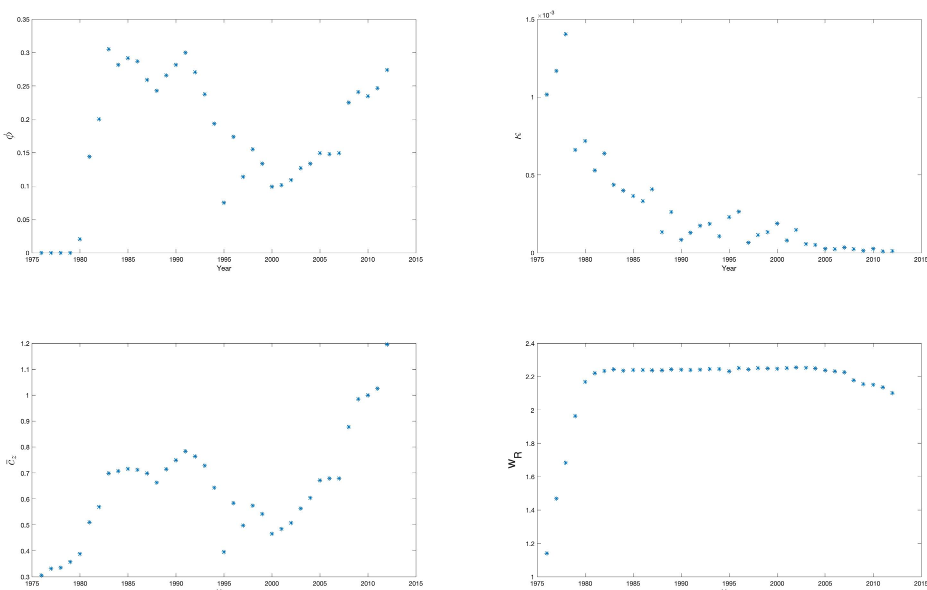


Figure 8: Model Parameters in Households and Firms

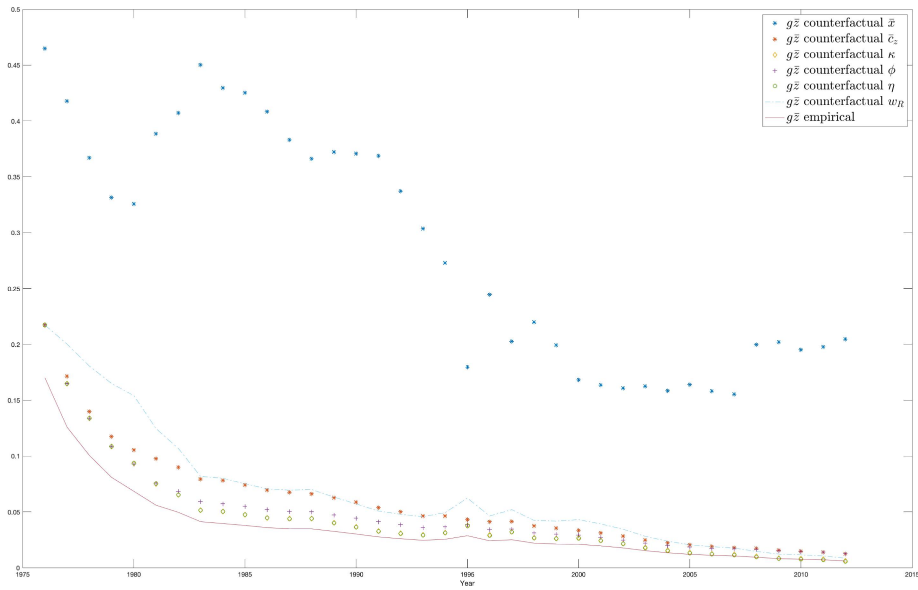


Figure 9: Counterfactual Economic Growth by Changing One Variable

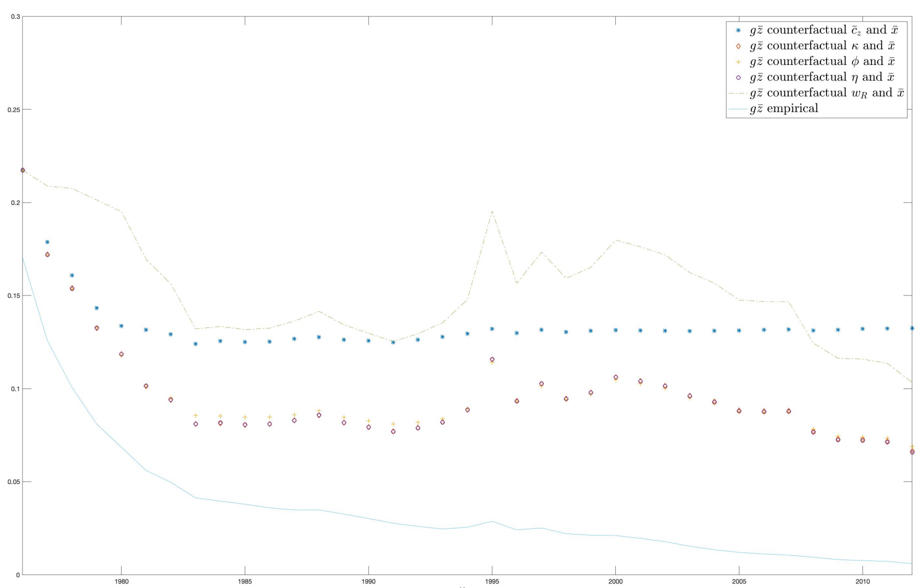


Figure 10: Counterfactual Economic Growth by Changing Two Variables

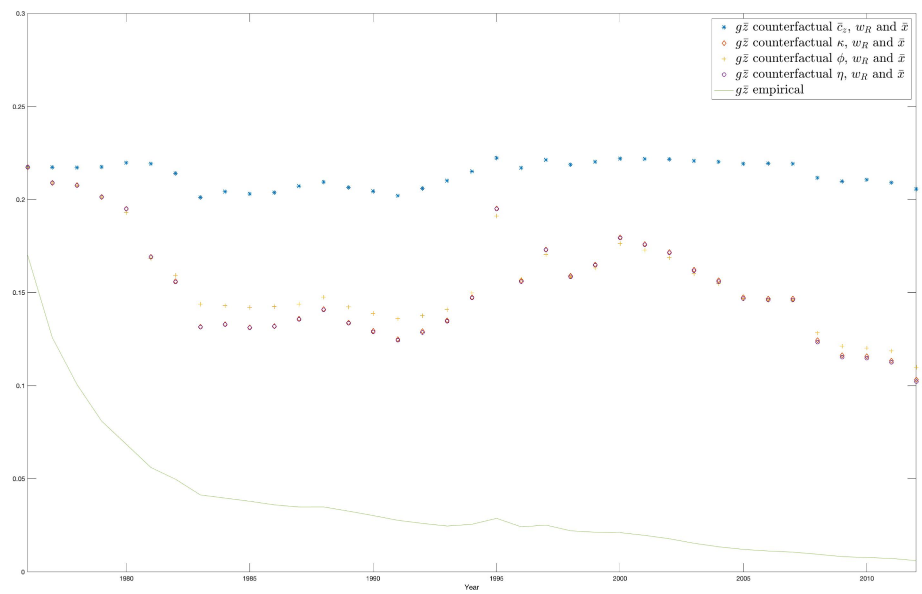


Figure 11: Counterfactual Economic Growth by Changing Three Variables