Technology Network, Innovation and Growth

Jingong (Jim) Huang *

University of Melbourne

This paper develops a multi-sector endogenous growth model which embeds a technology network that captures heterogeneous intersectoral knowledge spillovers. Each sector serves both as a distributor and a producer of knowledge. The interaction of these two forces influences long-run economic growth, sectoral shares and the firm size distribution. The sparsity of the network imposes an upper bound on the impact of knowledge spillovers. In this model, sectors converge to the same growth rate if they belong to the same *irreducible* network. However, their contributions to economic growth differ substantially, depending on their positions in the technology network and their efficiency in conducting innovation. Consequently, the model has implications for identifying key sectors in the economy. The gain in economic growth derived from promoting innovation in the sector that utilizes knowledge most efficiently is over 10,000 times larger than gain derived from promoting innovation in the least efficient sector.

Key words: Technology network, Firm innovation, Knowledge spillovers, Growth, Citation.

JEL Codes: O31 O33 O41 D85

*Department of Economics, University of Melbourne, 111 Barry Street, Carlton, VIC 3010, Australia. Email: jingongh1@unimelb.edu.au. I thank Matthew Greenwood-Nimmo for all the support and guidance. I also thank Nisvan Erkal, Mei Dong, Chris Edmond, Yves Zenou, Samuel Kortum, Hugo Hopenhayn, Bruce Preston, John Grigsby, Kei Kawakami, Lawrence Uren, Kevin Lee, Andy Dickerson, James Morley, May Li and the seminar participants at Melbourne University, 2017 SED, 2017 RES annual conference, 2016 North American Summer Meeting of the Econometric Society and 2016 Australasia Meeting of the Econometric Society for numerous comments and suggestions. All errors are my own.

1. INTRODUCTION

The seminal work by Klette and Kortum (2004) provides a parsimonious model that is able to explain several stylized facts about firm's innovation behavior. In their model, they treat firms as collections of knowledge and assume that past knowledge accumulation contributes to future innovation equally, regardless of which sector it comes from and which sector it contributes to. This simplifying assumption renders their model and the subsequent studies that have adopted the same framework silent on questions like how innovations in different sectors are connected with each other; what is the relationship between sectoral growth and aggregate economic growth; does industry-oriented policy have impacts beyond its targeted sector? These questions are of contemporary relevance as many governments, especially those of developing countries, have implemented ad hoc industrial policies without overall evaluations.[1] Unfortunately, little research has been done on these issues. In this paper, we aim to bridge the gap by building a framework incorporating a technology network that captures heterogeneous intersectoral technology dependence among sectors.

Figure 1 showcases an example of such a network constructed from patent citations.[2] In the graph, the economy is represented by a network. Each node is a sector in the economy and each edge implies the existence of a technology connection between sectors.[3] The thickness of an edge captures the strength of knowledge flows from one sector to another. The size of each node represents the strength of aggregate outward knowledge flows, measured by the cross-sector backward citation ratio.

Three features of Figure 1 are noteworthy. First, the strength of knowledge spillovers differs across sector pairs, as shown by the different thickness of the edges. Second, for a given sector pair, the knowledge spillover from one sector to the other is not the same as the spillover in the other direction, which implies that the network is asymmetric. Finally, knowledge spillovers do not exist among some sector pairs, which is reflected in the sparsity of the network. The above features all point to sector-pair-specific heterogeneity, one dimension of heterogeneity which is of interest in this paper.

[1] For example, the Chinese government proposed to invest more than 161 billion dollars over 10 years to develop the semiconductor industry in 2015.

[2] The patent citation data is from US Patent and Trademark Office and will be discussed in detail shortly. The figure is drawn using all citation data available in the dataset between 1975 and 2006 .

[3] The definition of sectors here follows Hall et al. (2001), who assign all patents to 37 sectors.

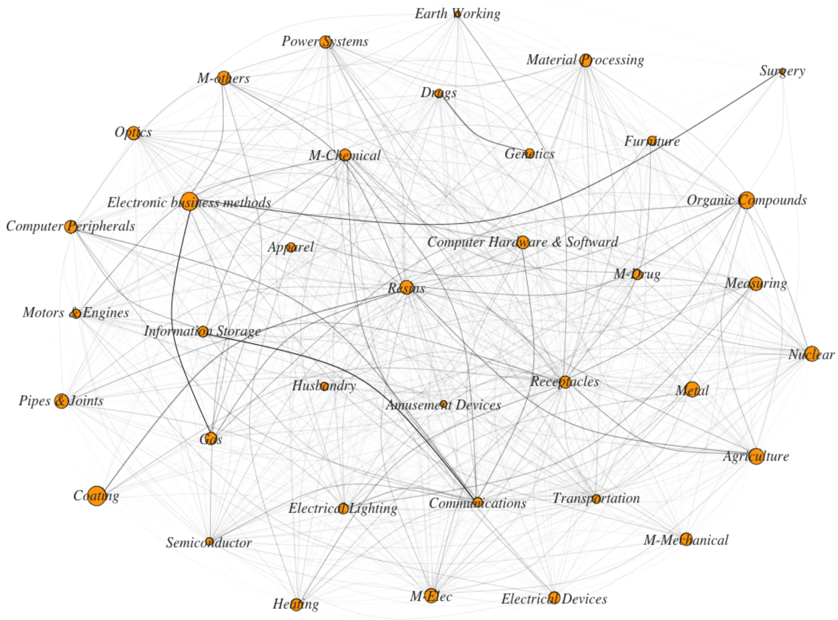


FIGURE 1

A TECHNOLOGY NETWORK REPRESENTATION OF ECONOMY

To further appreciate the magnitude of intersectoral knowledge flows, Figure 2 shows the kernel density of cross-sector citation ratios corresponding to the sectors in Figure 1. The horizontal axis measures the ratio of a sector's intersectoral citations over its total citations, while the vertical axis records the kernel density estimate, using a bandwidth of 0.02.[4] It is clear that cross-sector knowledge spillovers play a non-trivial role in the knowledge accumulation of all sectors, contributing from 17% to about 67% of knowledge flows with a mean of 39%.[5]

The importance of the technology network for innovation is substantiated by empirical evidence obtained from firm level data in this paper. First, intersectoral knowledge spillovers help firms to enter new sectors where they have a comparative technology advantage. We refer to this as the "pro-entry effect". Second, intersectoral knowledge spillovers are positively correlated with the accumulation of new patents, which we denote the "pro-innovation effect". The pro-innovation effect implies that a firm's knowledge accumulation in one sector will contribute to its innovation in another sector if there are knowledge spillovers from the former to the latter. These two empirical observations highlight both the extensive and intensive margins of knowledge spillovers. Last, we show that the pro-innovation effect is universal over all sectors, although the magnitudes differ substantially. This

[4]The intersectoral citations are calculated from all the other sectors to a given sector. The bandwidth is chosen to produce a smooth kernel curve.

[5]Note that these values are calculated for 2-digit sectors.

result indicates that firms in different sectors have heterogeneous efficiency in utilizing knowledge from other sectors.

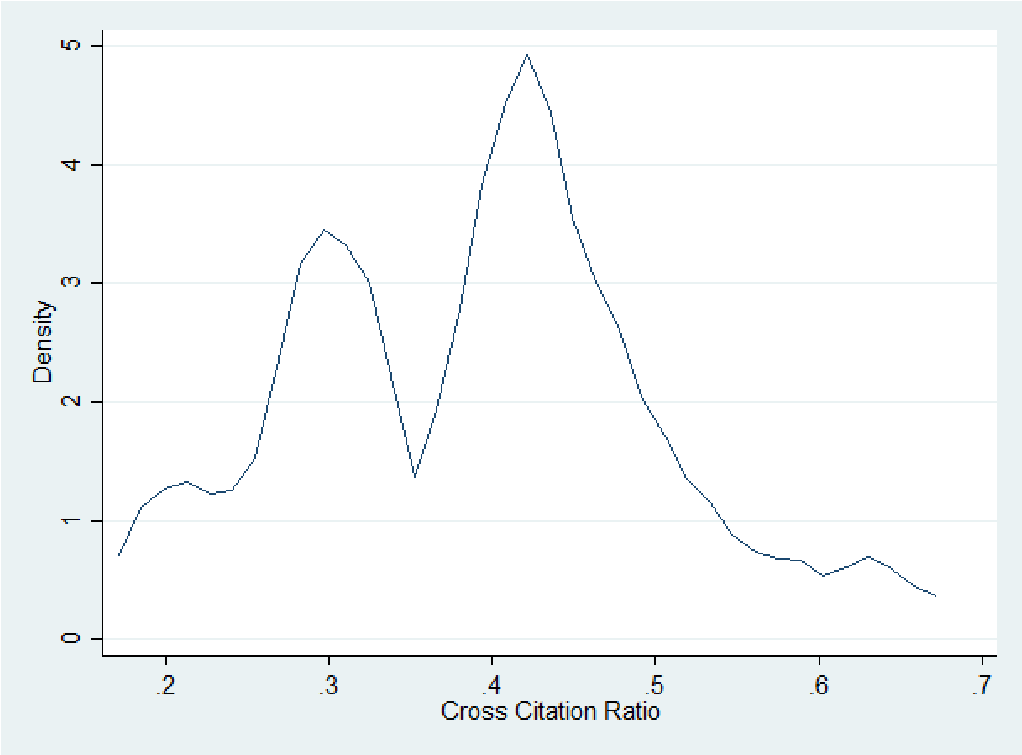


FIGURE 2

CROSS SECTOR CITATION RATIO

These empirical findings confirm the importance of intersectoral knowledge spillovers. However, one critical question remains unaddressed: how do the interactions between sectors affect growth at both the sector level and the aggregate level. To answer this question, we build a theoretical model, which allows firms to endogenously choose their innovation efforts across sectors, taking into account heterogeneous intersectoral knowledge spillover effects. The intersectoral knowledge spillover is an important feature, which determines the potential benefit of knowledge accumulation on cross-sector innovation. Motivated by the empirical analysis, we will assume that the technology network, which contains pair-wise knowledge spillovers for all sectors in the economy, is asymmetric and non-complete. The structure of the network is exogenous for individual firms.

Given the central role played by the technology network in our model, it is straightforward to note that different network structures have a significant impact on firms' cross-sector knowledge accumulation. To examine this impact, we study two types of networks in this paper: an irreducible network and a reducible network. The former represents an economy where each sector has knowledge spillovers to all the other sectors either directly or indirectly. In such an economy, knowledge spillovers display a global impact, which implies that knowledge accumulation in any sector will benefit the whole economy eventually. In contrast, an economy denoted by a reducible network can be thought of as a collection of multiple technology clusters. In this case, knowledge spillovers only exist among sectors within each technology cluster but not across clusters. The difference between the reducible and irreducible networks is determined by the sparseness of the network. When the network is sufficiently sparse, it becomes a reducible network.

In the benchmark case of the irreducible network, firms choose the sector-specific innovation rate, which depends on the average profitability of conducting innovation and the option value of research in a sector. Despite different innovation rates across sectors, in the long run, all sectors converge to the same growth rate due to the existence of global knowledge spillovers. However, different sectors contribute to economic growth differently, depending on their innovation capacities and their positions in the technology network to distribute knowledge. The interaction of these two forces determines long-run growth. In particular, the long-run growth rate is equal to the dominant eigenvalue of a matrix, Φ, that reflects the above interaction. A modified version of the matrix also determines the share of each sector. Specifically, we show that the share of a sector is represented by the sector's corresponding generalized eigenvector centrality associated with a matrix, $\Psi = \Phi + \mathrm{I}$,[6] which determines the dynamics of the sector share evolution. The generalized eigenvector centrality of a sector captures the sector's position in the downstream technology network.[7] A more central position implies higher knowledge spillovers from other sectors and thus a larger sector share. Lastly, the interaction of the intersectoral knowledge spillovers and heterogeneous sectoral innovation rates also plays a role in shaping the firm size distribution. The right tail of the firm size distribution is shown to be a Pareto distribution, and the thickness of the tail is pinned down by the ratio of the population growth rate and the long-run economic growth rate.

The previous theoretical results are established under the assumption that each sector displays global spillover effects. This assumption is relaxed when the network structure is assumed to be reducible. For a reducible network, the sectoral growth rate is driven by the local network structure of technology clusters that each sector belongs to and the converging behavior of sectoral growth is limited only within each technology cluster. As a result, different technology clusters demonstrate heterogeneous growth and aggregate economic growth is determined by the cluster that grows fastest. Given that sectors belong to different technology clusters may grow at different rates, their shares diverge in the long run. Moreover, the firm size distribution is also technology cluster specific. For

[6] I is an identity matrix.

[7] Eigenvector centrality is a measure of the importance of a node in a network. A higher score is assigned to nodes with connections to high score nodes. Generalized eigenvector centrality used in this paper takes into account the strength and the direction of each connection in addition to the number of connections.

firms that belong to different technology clusters, their size difference reflects the difference of those technology clusters' network structures.

We apply our framework to identify key sectors in the economy. A sector is a key sector if a small shock to the sector's innovation rate can lead to a large spillover effect. Many governments subsidize specific sectors to promote growth. Our analysis provides potential guidance on such policies from the perspective of knowledge spillovers. We find that there is substantial heterogeneity across sectors in terms of the induced impacts of sectoral shocks. Specifically, a 1% increase of the innovation rate for the most important sector can result in knowledge spillovers that are 10,000 times larger than in the case of the least important sector.

The remainder of this paper is organized as follows. Section 2 reviews the past literature. Section 3 presents the empirical motivation of this paper. The formal model is developed in Section 4, while policy analyses are provided in Section 5. Section 6 concludes.

2. Related Literature

This paper contributes to several strands of literature. Among others, it is closely related to the literature that studies R&D via endogenous innovation.[8] In particular, the model is an extension of Klette and Kortum (2004). In their paper, firms engage in a Schumpeterian-style innovation process and expand their products via a Poisson birth and death process. New products arise at a rate that depends on the knowledge accumulation embodied in past products, while some products of a firm are lost as a result of competition from rivals making them obsolete. A large literature has subsequently adapted their approach and applied it to different environments.[9] A common feature of these models is that innovation is undirected: that is, knowledge accumulation in one sector will benefit innovation in other sectors equally. In contrast, we construct a multi-sector model that focuses on the interdependency of sectors and allows for heterogeneous innovation rates. These new characteristics make it possible to think about the linkage between individual sectors and the aggregate economy and to evaluate the aggregate effects of industrial policies.

There is a large literature trying to model various networks existing in the economy. Among them, Jackson and Rogers (2007) are the first to present a general model of network formation in an attempt to explain the salient features of various networks formed in the real world. They argue that nodes in a network form links with each other either randomly or via local search using the existing

[8]See, for example,Romer (1986), Aghion and Howitt (1992), Aghion et al. (1997), Grossman and Helpman (1991) and Kortum (1997).

[9]See, for example, Lentz and Mortensen (2005, 2008), Acemoglu et al. (2013) and Akcigit and Kerr (ming)

network, the so-called meeting friends of friends. The technology network in this paper shares the fundamental feature of the social network in Jackson and Rogers (2007). Firms enter sectors where their existing patent portfolios have strong knowledge spillovers and innovate intensively in those sectors. This is similar to the local search in the social network.

Another strand of literature studies the size distribution of firms and cities. Luttmer (2007, 2011) describes a balanced growth model that features a Pareto distribution of firm size, which is consistent with the observed size distribution of U.S. firms' employment. In a different context, Gabaix (1999) employs a model, underlying which is a geometric Brownian motion with a reflecting barrier, to explain the city size distribution in the U.S.. The way in which these studies generate size differences is assuming that entities experience random growth and have different sizes because of the realization of different shocks.[10] This paper provides an alternative way to think about the firm size difference. In an economy represented by a sparse network, firms that belong to different technology clusters have different growth potentials, which are reflected by their sizes.

This paper is not the first attempt to incorporate networks into macro analysis. Oberfield (2012) develops a model of a business network through which firms form production chains and studies the endogenously emerging network. Cai and Li (2012) study the impact of intersectoral knowledge linkages on firms' innovation intensity and the sequence of entry into different industries. My paper treats the technology network as exogenous and explores how different structures of the network can affect economic growth, sectoral shares and the firm size distribution.

Finally, this paper contributes to the literature on knowledge spillovers. Jaffe (1986) constructs an empirical measure of technology spillovers to study the impact of the research of neighbouring firms on the success of a firm's R&D. He finds that high R&D firms tend to reap the benefit of knowledge spillovers while firms with low R&D are worse off. Bloom et al. (2013) study two types of spillovers: a positive technology spillover effect from other research firms and a negative business stealing effect from product market rivals. They employ and extend the measure of technology spillovers from Jaffe (1986) and conclude that the positive knowledge spillover effect dominates the negative business stealing effect.[11] Acemoglu et al. (2016) map the upstream technology network and sectoral patent growth to predict future innovation after 1995 and find strong predictive power at the sector level.

[10] Gabaix (2009) surveys the theory and application of the power law in economics and finance.

[11] There is a large amount of research in the IO literature that studies the external knowledge spillover between firms and their competition in the product market. A short list includes: D'Aspremont and Jacquemin (1988), Suzumura (1992),Amir and Wooders (1999),Anbarci et al. (2002) and Erkal and Piccinin (2010). My paper abstracts away from product market competition and external knowledge spillovers, and instead focuses on the internal knowledge spillovers. This feature makes my model tractable and renders the relationship between the network structure and the economic growth transparent.

My paper serves as the complement to the previous studies by providing novel empirical evidence of intersectoral knowledge spillovers at the firm level and new theoretical insights on the importance of intersectoral technological linkages.

3. EMPIRICAL EVIDENCE OF TECHNOLOGY NETWORK

In this section, we present evidence on the spillover effects of firms' past knowledge accumulation on future innovation. We ask three questions here: (i) Does a firm's existing patent portfolio affect its entry to a new sector in the future? (ii) Does a firm's current patent accumulation in a sector depends on its past patent accumulation in other sectors? and (iii) If so, how does this relationship vary across sectors? The first two questions present two margins of spillover effects: the former is what we term the pro-entry effect, which is intended to capture the extensive margin of knowledge spillover effects in expanding firms' patent portfolios. The latter is the pro-innovation effect. The pro-innovation effect captures whether past knowledge accumulation in one sector helps to promote innovation in technologically related sectors, representing an intensive margin of knowledge spillover effects. The third question deals with the heterogeneity of the efficiency of different sectors in absorbing and applying intersectoral knowledge.

The backbone for our analysis here is the NBER Patent Database from the United States Patent and Trademark Office (USPTO) from 1975 to 2006. The dataset contains patent records that provide information on the unique assignee number of the inventor of each patent, the country the assignees belong to, the date each patent is applied for and granted as well as the technology sector each patent belongs to, etc. We use the field classification proposed by Hall et al. (2001), who assign all patents into 37 technology sectors. Details of how each sector is defined can be found in Hall et al. (2001). Here, for convenience, we label each technology sector by a number from 1 to 37. TABLE A.1 in the Appendix shows the correspondence between the numeric label and the original field. We only use patents granted to U.S. companies and drop observations for which the assignees are missing. The patent database contains an associated dataset upon the citing-cited relationship of patents. The citation data will be critical for us to empirically construct the technology network that captures the strength of knowledge flows across sectors. Estimations in this paper will be based on the data between 1990 and 2001. The 2001 end date is chosen to allow for a 5-year window for patent reviews. The 1990 start date allows enough pre-sample data to implement the empirical strategy that will be discussed in detail later. Estimations extending the range of data deliver qualitatively same results.

To address the first question, some idea about how close two technologies are, or more precisely, how easily certain technology can be applied to others is required. Therefore, we introduce a measure of technology applicability based on the patent citation data. A citation may serve as an indicator of knowledge spillovers. For instance, if patent A cites patent B, then the implication is that knowledge flows from B to A. My measure of technology applicability from sector y to x, $g(y, x)$, is constructed using the following formula:

$$g(y, x) = \frac{N(\text{citation from } x \text{ to } y)}{N(\text{total citation to } y)}. \tag{1}$$

This measure is sector-pair specific and can be interpreted as the average spillover effect from sector y to sector x. Using this measure, we can continue to construct the average proximity between a firm's current patent portfolio and any target technology sector. The measure of average proximity is needed because firms can operate in multiple sectors, and our first measure cannot account for this. We use the variable $Mproximity^x_{f,t}$ to denote the average proximity, and it is calculated as

$$\sum_{y \in \{y \neq x\}} \frac{n^y_{f,t}}{n_{f,t}} g(y, x),$$

where $n^y_{f,t}$ is the patent stock for firm f at time t in sector y and $n_{f,t}$ is total patent stock at t.[12] Note that the target sector x is excluded in the summation because we want to analyze the intersectoral spillover effects. With the above measure, we run a Probit regression over a firm's probability of entering a given sector on its average proximity to this sector:

$$Pr(y^x_{f,t}) = \alpha_1 Mproximity^x_{f,t-1} + \alpha_2 Controls^x_{f,t-1} + \mu^x_f + \vartheta_t + \zeta^x + \varepsilon^x_{f,t}, \tag{2}$$

where the dependent variable $y^x_{f,t}$ is a dummy variable that takes the value of 1 if the corresponding firm f produces patents in technology sector x at time t and zero otherwise. We control for whether firm f has previously operated in the same sector, $y^x_{f,t-1}$, as well as the total number of technology sectors firm f operates in at $t-1$, $N_{f,t-1}$. To deal with unobserved heterogeneity, it is assumed that the error term is composed of μ^x_f, a firm-sector fixed effect, a full set of year dummies, ϑ_t, sector dummies, ζ^x, and an idiosyncratic component, $\varepsilon^x_{f,t}$. We follow Wooldridge (2002) and instrument the unobserved firm-sector fixed effect μ^x_f with the time mean of all exogenous variables and the initial value of the dependent variable y^x_0. The key parameter that we are interested in is α_1, which captures the pro-entry effect. A positive α_1 implies that a patent portfolio that is more technically related to sector x contributes to the firm's entry into sector x.

Estimation results are presented in TABLE 1. As shown in column (1), a firm is more likely to enter a technology sector x when its patent portfolio contains technologies closer to that sector. The estimates remain significant after controlling for various factors (column (2) to (4)). The additional

[12] $n^y_{f,t}$ can be zero if firm f does not hold any patent in sector y at time t.

control variables behave as we would expect. In particular, a firm that currently owns patents in sector x is more likely to continue innovation in the same sector (column(2)), and there is a higher probability that a firm with a broad set of patents will expand into another sector (column(3)). The magnitude and significance of the above variables are largely unaffected by adding time and sector dummies to the regression as shown in column (4). The regression results taking account of unobserved firm-sector fixed effects are provided in column (5). After controlling for the unobserved firm-sector fixed effects, the coefficients of all controls decline significantly with $Mproximity^x_{f,t}$ being the only exception. In fact, the estimate for $Mproximity^x_{f,t}$ increases compared to its estimate in column (4), demonstrating the robustness of our main results.

TABLE 1

PROBIT REGRESSIONS

	(1)	(2)	(3)	(4)	(5)
$Mproximity^x_{f,t-1}$	14.75***	13.55***	14.16***	11.89***	14.57***
	(0.17)	(0.12)	(0.13)	(0.14)	(0.24)
$y^x_{f,t-1}$		2.17***	1.98***	1.90***	1.62***
		(0.01)	(0.01)	(0.01)	(0.01)
$N_{f,t-1}$			0.06***	0.06***	0.00
			(0.00)	(0.00)	(0.00)
$YearDummy$	No	No	No	Yes	Yes
$SectorDummy$	No	No	No	Yes	Yes
$Firm\text{-}sector\ FE$	No	No	No	No	Yes
$Constant$	-1.66***	-2.08***	-2.25***	-2.65***	-2.80***
	(0.01)	(0.01)	(0.00)	(0.02)	(0.03)
Observations	2,042,178	2,042,178	2,042,178	2,042,178	2,042,178

A */**/*** next to the coefficient indicates significance at the 10/5/1% level. The variable . $Mproximity_{y,t-1}$ is constructed using the data at time $t-1$. Firm-sector fixed effect is instrumented using the method proposed by Wooldridge (2002). All standard errors are clustered at firm level.

One concern with the above approach is that the positive relationship between technology proximity and the probability of entering new sectors may be driven by a few sectors that demonstrate strong positive effects instead of a universal phenomenon across all sectors. To address such a concern, TABLE A.2 in the Appendix reports the results of Probit regressions for each technology sector.

The conclusion drawn there is that the pro-entry effect is indeed observed in all sectors, though there is some heterogeneity regarding how strong the effect is across sectors. In general, the probability of entering a technology sector in the future is higher if a firm currently owns a patent portfolio that is technically closer to the sector of interest.

We now turn to evaluate how much past knowledge accumulation contributes to the building of new knowledge. We construct a new measure, $Wpatent^{x}_{f,t}$, to capture the weighted patent stock with the weights equal to $g(y,x)$. $Wpatent^{x}_{f,t}$ for firm f at time t with respect to sector x is defines as

$$\sum_{y\in\{y\neq x\}} I(y_{f,t}=1)g(y,x)n^{y}_{f,t},$$

where $n^{y}_{f,t}$ is the patent stock for firm f in the technology class y at time t and $I(y_{f,t}=1)$ is a dummy variable that takes the value of 1 if firm f does not hold a patent in sector y, and 0 otherwise. This measure captures the intersectoral spillover effect of the current patent portfolio with respect to sector x. Note that we do not include the patent stock in x. We follow Hall et al. (2005) and construct the patent stock using a perpetual inventory method with a 15% depreciation rate.[13] That is, $n^{y}_{f,t}=\Delta n^{y}_{f,t}+(1-\delta)n^{y}_{f,t-1}$, where $\Delta n^{y}_{f,t}$ is the number of new patents firm f produces in sector y at time t and $\delta=0.15$. We then run the following Negative Binomial model:

$$PatentCounts^{x}_{f,t}=\beta_1 Wpatent^{x}_{f,t-1}+\beta_2\, Controls^{x}_{f,t-1}+\tilde{\mu}^{x}_{f}+\tilde{\vartheta}_t+\tilde{\zeta}^{x}+\tilde{\varepsilon}^{x}_{f,t} \qquad (3)$$

where $PatentCounts^{x}_{f,t}$ represents the number of new patents produced by firm f in technology sector x from time $t-1$ to time t and $\tilde{\mu}^{x}_{f}$ is the firm-sector fixed effect. $\tilde{\vartheta}_t$ and $\tilde{\zeta}^{x}$ are sets of time dummies and sector dummies respectively. Due to the nonlinearity of the Negative Binomial model, we follow Blundell et al. (1999) and use the pre-sample mean scaling method to control for fixed effects. The idea is to use pre-sample data on patenting behavior to instrument for unobserved heterogeneity. The long panel of patent data from USPTO allows me to construct the pre-sample average between 1970 and 1989. The controls used here are the same as those in (2) except that we include a dummy to indicate whether a firm previously innovates in the target sector and the lag value of patent counts of the firm in the target sector.

The coefficient of interest is β_1. A positive sign for this coefficient implies the existence of a pro-innovation effect. The estimates are shown in TABLE 2. From column (1), it is clear that a firm's production of new patents is positively related to the firm's past patent accumulation in other sectors. Adding variables that control for whether firm f owns a patent and how many patents firm f owns decreases the coefficient for $Wpatent^{x}_{f,t}$ (column 2). Nonetheless, the inclusion of further controls

[13] Hall et al. (2005) use R&D to calculate knowledge capital while we use patents here.

has little effect on it (column (3) to (4)). Finally, in column (5), we demonstrate the results using the pre-sample mean scaling method to control for the firm-sector fixed effect, which changes neither the magnitude nor the significance of the estimate for the weighted patent stock variable. As shown in column (5), past knowledge accumulation has a substantial effect on the innovation of new patents. A 1% increase of the weighted patent stock leads to a 0.4% increase in the production of new patents.

TABLE 2

NEGATIVE BINOMIAL REGRESSIONS

	(1)	(2)	(3)	(4)	(5)
$Wpatent^x_{f,t-1}$	0.76***	0.40***	0.43***	0.39***	0.39***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$I(x_{f,t-1} = 1)$		2.87***	2.88***	2.82***	2.66***
		(0.04)	(0.04)	(0.04)	(0.04)
$PatentCounts^x_{f,t-1}$		0.06***	0.06***	0.06***	0.04***
		(0.01)	(0.01)	(0.01)	(0.01)
$N_{f,t-1}$			-0.02***	-0.01**	-0.00
			(0.00)	(0.00)	(0.00)
$Year Dummy$	No	No	No	Yes	Yes
$Sector Dummy$	No	No	No	Yes	Yes
$Firm\text{-}sector\ FE$	No	No	No	No	Yes
$Constant$	-0.47***	-2.22***	-2.09***	-2.80***	-2.73***
	(0.02)	(0.02)	(0.02)	(0.07)	(0.09)
$ln\alpha$	2.20***	0.73***	0.72***	0.72***	0.55***
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Observations	2,029,470	2,029,396	2,029,396	2,029,396	1,050,316

A */**/*** next to the coefficient indicates significance at the 10/5/1% level. $I(x_{f,t-1} = 1)$ is the indicator function that takes the value of 1 if firm f operates in sector x at time $t-1$ and 0 otherwise. $Wpatent^x_{f,t-1}$ is in log form to make the interpretation more convenient. The firm-sector fixed effect is instrumented using the method proposed by Blundell et al. (1999) and is reported in column (5). All standard errors are clustered at the firm level.

The above analysis studies the average effect of intersectoral knowledge spillovers. In the rest of this section, the focus will be on the heterogeneity of intersectoral knowledge spillovers. To this end, we run the same regression as (3) but restrict the sample to each sector instead of pooling all sectors together. The estimate β_1 for each sector can be interpreted as the average efficiency of firms in absorbing knowledge from other sectors to their own sectors.

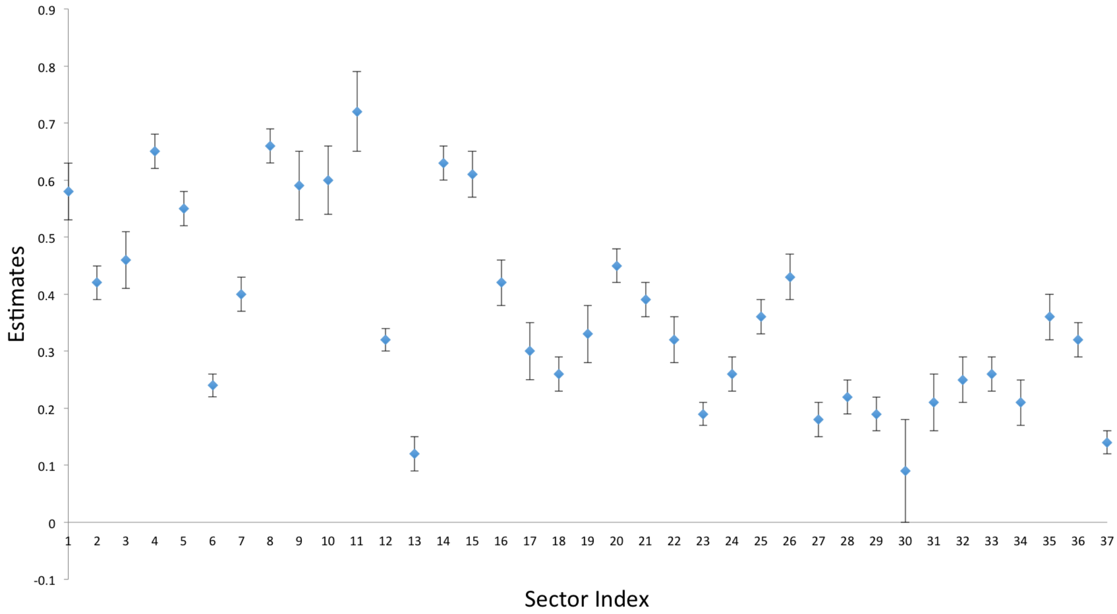


FIGURE 3

ESTIMATES OF SECTOR INNOVATION INTENSITIES

Each sector's estimate with one standard error is displayed in FIGURE 3.[14] There is significant heterogeneity across sectors with the parameter estimates, ranging from 0.09 to 0.72 with a mean of approximately 0.3. The results reveal that intersectoral knowledge spillovers contribute significantly to knowledge building for most sectors. However, each sector's ability to absorb and use knowledge to produce research output differs substantially.

To sum up, this section presents empirical evidence about the importance of intersectoral knowledge spillovers on innovation at the firm level. In the next section, we will build a model that embeds the features shown here. In particular, the model will include two layers of heterogeneity: intersectoral knowledge spillovers and the sector-specific innovation rate. The latter endogenously arises as the option value of research is different across sectors.

4. THEORETICAL MODEL

4.1. Multi-sector Innovation Model with an Irreducible Technology Network

There are a finite number of technology sectors in the economy. Each sector is a collection of technologies that share similar features. Sectors differ in their technology spillovers to other sectors.

[14]Details of the regression results are shown in TABLE A.3 in the Appendix.

The heterogeneous sector-pair technology spillovers (edges) along with all sectors (nodes) form a technology network. Formally, the technology network is modeled as a $M \times M$ weighted directed adjacency matrix G, where each entry $g_{ij} \geq 0$ denotes the knowledge flow from sector i to sector j, and M is the number of sectors. There are intersectoral technology spillovers from i to j if $g_{ij} > 0$. The implications of different network structures will be explored here. As a benchmark case, we will study an irreducible technology network, whose definition is formalized as follows:

Definition 1. A network is irreducible if $\forall i, j$, there exists a sequence of indices $l_1, l_2, \cdots, l_n$ such that $g_{il_1} g_{l_1 l_2} \cdots g_{l_n j} > 0$.

Loosely speaking, an irreducible technology network represents an economy where all sectors have potential impacts on each other either directly or indirectly.

Entrepreneurs develop their knowledge portfolios based on their past knowledge accumulation. The innovation activity is modeled as a Poisson process. A novel feature is that old knowledge not only contributes to the building of new knowledge in the same sector but also in other sectors. In other words, innovation in one sector demonstrates both intrasectoral and intersectoral knowledge spillover effects. The strength of the intersectoral knowledge spillover effects is captured by the off-diagonal elements of the adjacency matrix G.

Incumbents use patents to store new knowledge. Each patent possesses some value to the owner. We will assume that the value of patents is randomly distributed with a sector-specific mean, $\bar{\pi}_i$. In general, the average value of a patent can be different across sectors, meaning that $\bar{\pi}_i \neq \bar{\pi}_j$ for $i \neq j$. There are several ways to interpret the value of a patent. A firm can use its patent to design a new product, and thus enjoy the monopoly profit of the product. Alternatively, a new patent may add value to a firm's current product and improve its quality. In an environment where quality matters to consumers, the two can be equivalent. Here, we take the first interpretation. Now assume that a firm at time t has n_i patents in the technology sector i, and the associated values of these patents are denoted by a vector $\tilde{\pi}_i = (\pi_{i1}, \pi_{i2}, \cdots, \pi_{in_i})$. In addition, for a firm with a total of n patents, the value of those patents is represented by a vector of vectors $\tilde{\pi}^n = (\{\tilde{\pi}_i\}_{i \in \mathcal{T}})$, where $n = \sum_{i \in \mathcal{T}} n_i$ and $\mathcal{T}$ is the set of all sectors. Note that if a firm does not have any patent in a certain sector s, then $n_s = 0$.

Each firm can be regarded as a collection of research teams, each specializing in a certain technology sector. Each research team pools knowledge of other teams together and devotes effort in order to produce new patents. However, knowledge from different sectors will not contribute symmetrically to innovation. In particular, it is assumed that 1 unit of knowledge stock (1 patent) in sector j will serve as g_{ji} units of effective knowledge stock when used for innovation in sector i. Therefore, the total effective knowledge stock in sector i for a firm with the patent portfolio specified above is $\sum_{j\in\mathcal{T}} n_j g_{ji}$. This term can be interpreted as total knowledge spillovers to sector i.

All firms have access to a common production technology that allows them to innovate at the rate λ_i per patent in sector i. The cost for a research team specializing in sector I is assumed to be a function of the total innovation rate and the effective knowledge stock, and takes the following form:

$$C\Big(\lambda_i \sum_{j\in\mathcal{T}} n_j g_{ji}, \sum_{j\in\mathcal{T}} n_j g_{ji}\Big).$$

$C(.,.)$ is assumed to be homogeneous of degree one and increases with both arguments. As a result, the cost function can be written as:

$$C = \sum_{j\in\mathcal{T}} n_j g_{ji} c(\lambda_i),$$

where the variable cost function is specified as follows:

$$c(\lambda_i) = \lambda_i^{\epsilon}\theta^{1-\epsilon},\ \epsilon > 1. \tag{4}$$

θ is a technology parameter which is constant across sectors. The assumption $\epsilon > 1$ ensures that the cost function is a convex function of the flow innovation rate. In addition, $c(.)$ is assumed to be twice differentiable.

Given the previous setup, a firm takes as given its current portfolio of patents and decides the optimal innovation rate for each sector by solving the following value function:

$$\begin{aligned} rV(\tilde{\pi}^n) - \dot{V}(\tilde{\pi}^n) = \max_{\{\lambda_i\}_{i\in\mathcal{T}}} \Big\{ &\sum_{i\in\mathcal{T}} \sum_{s=1}^{n_i} \pi_{is} - \sum_{i\in\mathcal{T}} c(\lambda_i) \sum_{j\in\mathcal{T}} n_j g_{ji} \\ &+ \sum_{i\in\mathcal{T}} \lambda_i \sum_{j\in\mathcal{T}} n_j g_{ji} (E[V(\tilde{\pi}^{n+1})|i] - V(\tilde{\pi}_n)), \end{aligned} \tag{5}$$

where $E[V(\tilde{\pi}^{n+1})|i] = E[V(\tilde{\pi}^n \cup \pi_{i(n_i+1)})]$ and r is the interest rate.

The first line on the right-hand side captures net flow profits generated by the firm's current patent portfolio over the total cost incurred for innovation. The second line represents the expected change of the firm's value resulting from the arrival of new patents across sectors. The expected value change is the sum of all sectoral changes. For a given sector, say i, in which the firm currently operate, $\lambda_i \sum_{j\in\mathcal{T}} n_j g_{ji}$ is the sector-specific aggregate innovation rate, while $E[V(\tilde{\pi}^{n+1})|i] - V(\tilde{\pi}_n)$ is the expected increase in a firm's value conditional on the arrival of a new patent in sector i.

The next proposition characterizes the solution to the value function and the associated optimal innovation choices.

Proposition 1. *The value function for the firm's problem has a solution as follows:*

$$V(\tilde{\pi}^n) = \sum_{i\in\mathcal{T}} \sum_{s=1}^{n_i} \frac{\pi_{is}}{r} + \sum_{j\in\mathcal{T}} n_j R_j, \tag{6}$$

where $(R_1 \cdots R_M)$ *are the solutions to a system of* M *nonlinear equations:*

$$rR_j = \sum_i \theta(\frac{\frac{\bar{\pi}_i}{r} + R_i}{\epsilon})^{\frac{1}{\epsilon-1}} g_{ji}[(1 - \frac{1}{\epsilon})(\frac{\bar{\pi}_i}{r} + R_i)] \quad \forall j \in \mathcal{T}, \tag{7}$$

and the optimal innovation choices are given by:

$$\lambda_i = \theta(\frac{\frac{\bar{\pi}_i}{r} + R_i}{\epsilon})^{\frac{1}{\epsilon-1}}. \tag{8}$$

Proof. See Appendix. □

From the proposition, we can see that the value function of a firm with a patent portfolio $\tilde{\pi}^n$ is equal to the discounted additive sum of random profits generated from the firm's whole portfolio plus the aggregation of the sector-specific option value of research, $\sum_i n_i R_i$. It is interesting to note that the option value of research in each sector is itself a function of the option value of research of all the other sectors. This is an intrinsic feature of the model because each piece of knowledge created in one sector will be useful for innovation in other sectors. They are interconnected with each other through the channel of intersectoral knowledge spillovers, which is captured by the off-diagonal elements in G. If we shut down the channel of intersectoral knowledge spillovers, the option value of research for different sectors becomes independent of each other.

The optimal innovation rate for each sector depends on both the average profit and the research value of patents in the sector. A higher average profit or a higher option value of research induces firms to increase their effort and thus results in a higher innovation rate.

To complete our analysis, it is necessary to specify the entry of new firms. Suppose that the population is growing at a constant rate η in the economy. There is a fixed proportion of each generation that has the potential to become entrepreneurs. The proportion of potential entrepreneurs is constant over time. As a result, the number of firms grows at the same rate as the population. Further assume that for a firm to obtain an entry rate of 1, it must pay a fixed cost f. We specify the following free entry condition:

$$E[V(\pi)] = f,$$

where $E[V(\pi)]$ is the expected value of an entrant. Upon entry, firms draw from a common probability distribution function that determines which sector they will enter. Specifically, let p_i denote the probability of entering sector i. The expected value of an entrant is therefore:

$$E[V(\pi)] = \sum_i p_i V(\pi_i).$$

The probability distribution function of entry is assumed to be exogenous. This assumption is innocuous for the main results in this paper.

4.2. Dynamics of Firms' Innovation

As shown in the previous section, for a firm with patent portfolio $\{n_{i,t}\}_{i\in\mathcal{T}}$, the increase of the patent stock in sector i for a small time interval Δt is:

$$n_{i,t+\Delta t} - n_{i,t} = \lambda_i \Delta t \sum_{j\in\mathcal{T}} n_{j,t}\, g_{ji} \;\; \forall i \in \mathcal{T} \tag{9}$$

where $n_t^i - n_{t+\Delta t}^i$ is the accumulation of new patents in technology sector i for the time interval Δt. The above equation reveals that the evolution of a firm's patent portfolio is a nested function of the technology network and sectoral innovation rates. To see this, rewrite equation (9) in matrix form:

$$\begin{bmatrix} n_{1,t+\Delta t} - n_{1,t} \\ n_{2,t+\Delta t} - n_{2,t} \\ \vdots \\ n_{M,t+\Delta t} - n_{M,t} \end{bmatrix} = \begin{bmatrix} \lambda_1 g_{11} & \lambda_1 g_{21} & \cdots & \lambda_1 g_{M1} \\ \lambda_2 g_{12} & \lambda_2 g_{22} & \cdots & \lambda_2 g_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_M g_{1M} & \lambda_M g_{2M} & \cdots & \lambda_M g_{MM} \end{bmatrix} \begin{bmatrix} n_{1,t} \\ n_{2,t} \\ \vdots \\ n_{M,t} \end{bmatrix} \Delta t$$

or more compactly in a continuous form:

$$\dot{n}_t = \Phi n_t$$

where

$$\dot{n}_t = \begin{bmatrix} \dot{n}_{1,t} \\ \vdots \\ \dot{n}_{M,t} \end{bmatrix} \quad \Phi = \begin{bmatrix} \lambda_1 g_{11} & \lambda_1 g_{21} & \cdots & \lambda_1 g_{M1} \\ \lambda_2 g_{12} & \lambda_2 g_{22} & \cdots & \lambda_2 g_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_M g_{1M} & \lambda_M g_{2M} & \cdots & \lambda_M g_{MM} \end{bmatrix}$$

Matrix Φ captures a modified technology network which is adjusted by sectoral innovation rates. A firm's patent accumulation in the past affects its future patent accumulation through knowledge spillovers, the strength of which is captured by matrix Φ. A higher innovation rate in the i_{th} sector leads to faster accumulation of patents in sector i and also contributes to patent accumulation in sector j since $n_{i,t}g_{ij}$ enters the expression for $n_{j,t+\Delta t} - n_{j,t}$ given that $g_{ij} \neq 0$. Even if $g_{ij} = 0$, as long as the network is irreducible, then there exists a path $(i, l_1, l_2, \cdots, l_n, j)$ so that $n_{i,t}$ impacts $n_{j,t+\Delta t} - n_{j,t}$ indirectly. The existence of an irreducible network guarantees that each sector *communicates* with each other.

It is interesting to compare the technology network here and the social development network in Jackson and Rogers (2007). In their paper, the main way for a person to make new friends is by meeting friends of friends. In particular, if a person has a well-connected friend, it will be much easier for the person to make new friends through the local search of her well-connected friend's social network. In our case, new firms enter a certain sector of the economy and accumulate knowledge in that sector. They are more likely to subsequently enter new sectors where they can apply their existing knowledge more efficiently. This is analogous to the local search in the social networks.

Due to the fact that the arrival of new patents is random, the evolution of a firm's patent portfolio is history dependent which makes it a daunting job to track the dynamics of patent accumulation across sectors. In order to proceed with the analytical analysis, the rest of the paper adopts the mean-field approximation, popularized by Jackson and Rogers (2007). The mean-field approximation assumes that all innovations happen deterministically at the expected rate. Under the mean-field approximation, aggregating individual firms' patent accumulation and taking into account entrants' innovation, we obtain the sectoral patent accumulation as follows:

$$N_{i,t+\Delta t} - N_{i,t} = p_i L_t \eta \Delta t + \lambda_i \Delta t \sum_{j \in \mathcal{T}} N_{j,t} g_{ji} \;\; \forall i \in \mathcal{T}, \tag{10}$$

where $N_{i,t}$ is the total number of patents in sector i, L_t is the population of firms at time t and p_i is the probability that a new entrant will enter sector i. The increase of the number of patents in sector i for a short time interval Δt is the result of both entrants' and incumbents' innovation. We are free to aggregate individual firms' knowledge stocks, $n_{i,t}$, to the sectoral level because of the constant return to scale of the innovation function. The next proposition characterizes growth at both the aggregate and sectoral level, where growth is defined as the increase in the stock of knowledge.

Proposition 2. *In the long run, different sectors converge to the same growth rate, which is equal to the aggregate growth rate of the economy:*

$$\dot{N}_t/N_t = \dot{N}_{i,t}/N_{i,t} = \tau^* \textit{ if } \eta < \tau^*$$
$$\dot{N}_t/N_t = \dot{N}_{i,t}/N_{i,t} = \eta \textit{ if } \eta > \tau^*$$

where τ^* is the dominant eigenvalue of the matrix Φ.

Proof. See Appendix. □

In an economy where each sector is connected with each other, it is no surprise that all sectors eventually grow at the same rate. There are two sources of growth in the economy: both entrants and incumbents contribute to the accumulation of knowledge. On one hand, the dominant eigenvalue of Φ determines how fast knowledge in every sector grows due to incumbents' innovation. On the other hand, entrants in every period bring new ideas to the economy at the rate of population growth. Ultimately, the magnitude of economic growth depends on whether the former dominates the latter or not. The rest of the paper will focus on the theoretical interesting case of $\tau^* > \eta$, namely, the incumbents' contribution to innovation outweighs the entrants' contribution.

Proposition 2 sheds light on our understanding of growth in an economy where intersectoral knowledge spillovers prevail. It highlights the role of the innovation-adjusted technology network in determining economic growth. For an economy without intersectoral knowledge spillovers, economic growth is simply determined by the fastest growing sector. Here every sector contributes to the knowledge accumulation in the economy in two ways. First, each sector applies both intrasectoral and intersectoral knowledge to conduct innovation. Their innovation capacities determine how much knowledge they can produce. Second, knowledge produced by each sector can be used for innovation by other sectors. The position of a sector in the technology network determines how far this sector's knowledge spreads in the economy. It is the interplay of these two forces that determines the growth rate of the economy.

Given the *equal* growth rate of all sectors in our economy, it naturally gives rise to a stationary patent distribution over sectors. To characterize this feature, we first define a vector $Q_t = (Q_{1,t}, Q_{2,t}, \cdots, Q_{M,t})$, where $Q_{i,t} = \frac{N_{i,t}}{N_t}$ is sector i's share of total patent stock at time t. Let $\Delta t = 1$, and rewrite equation (10) as follows:

$$N_{i,t+1} = p_i L_t \eta + N_{i,t} + \lambda_i \sum_{j \in \mathcal{T}} N_{j,t} g_{ji} \;\; \forall i \in \mathcal{T} \tag{11}$$

Define

$$\Psi = \begin{bmatrix} 1+\lambda_1 g_{11} & \lambda_1 g_{21} & \cdots & \lambda_1 g_{M1} \\ \lambda_2 g_{12} & 1+\lambda_2 g_{22} & \cdots & \lambda_2 g_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_M g_{1M} & \cdots & \cdots & 1+\lambda_M g_{MM} \end{bmatrix}$$

Now equation (11) can be written in matrix form:

$$\tilde{N}_{t+1} = L_t \eta \tilde{p} + \Psi \tilde{N}_t \tag{12}$$

where $\tilde{p} = (p_1, p_2, \cdots, p_M)$ and $\tilde{N}_t = (N_{1,t}, N_{2,t}, \cdots, N_{M,t})'$. Matrix Ψ determines the evolution of each sector's share. The next proposition summarizes the main findings.

Proposition 3. *The sector shares Q_t approach to a limit Q in the long run, which satisfy the following system of equations:*

$$\psi^* Q = \Psi Q$$

where $\psi^* = \tau^* + 1$.

Proof. See Appendix. □

The above proposition provides new insights on the share of each sector. Note that Q is the eigenvector associated with the dominant eigenvalue ψ^* of the matrix Ψ. The share of each sector is equal to the corresponding entry of the eigenvector. In the network literature, Q is called the *generalized eigenvector centrality*. This is a measure of the importance of nodes in the network. It assigns scores to each node. A higher score implies a more central position of a node. A node gets higher scores if it is connected with other high score nodes. In the context of this paper, a sector's share in the long run is determined by its position in the network, represented by the matrix Ψ. A sector that receives strong knowledge spillovers from other sectors is located in a central position of the network.

Note that a sector's share here is very different from that in a model without intersectoral knowledge spillovers. For the latter, the largest sector is always the one with the highest innovation rate. In contrast, with intersectoral knowledge spillovers, that may not be the case. A sector that innovates at a slow rate may turn out to have a large size if it enjoys strong knowledge spillovers from other sectors. The following example illustrates this point. Suppose there are 4 sectors in the economy, and the network structure of the economy is represented by a weighted adjacency matrix as follows:

$$G' = \begin{bmatrix} 1 & 0.3 & 0.3 & 0.3 \\ 0.1 & 1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 1 \end{bmatrix}$$

The network structure in the above matrix is such that sector 1 receives the strongest knowledge spillovers from other sectors in the economy. Now let's assume that the innovation rates for each sector are:

$$\lambda_1 = 0.01,\ \lambda_2 = 0.011,\ \lambda_3 = 0.012,\ \lambda_4 = 0.013,$$

where sector 1 has the lowest innovation rate. Given this information, the long-run sector shares can be easily calculated according to Proposition 3:

$$Q_1 = 0.3236,\ Q_2 = 0.1727,\ Q_3 = 0.2193,\ Q_4 = 0.2844.$$

From the above results, sector 1 is the largest sector in the economy. This simple numerical example showcases a possibility that a sector can become the dominant sector even if it does not innovate as fast as the other sectors do. Sectors that are capable of using resources from other sectors efficiently can grow large.

We have shown how heterogeneous intersectoral knowledge spillovers together with sectoral innovation rates affect growth and sector shares. Next, we will demonstrate that the same forces also play a role in affecting the size distribution of firms.

The size of a firm is measured by the number of patents it holds. Aggregating equation (9) across all sectors gives the dynamics of firm size:

$$n_{t+\Delta t} - n_t = \sum_{i \in \mathcal{T}} \lambda_i \Delta t \sum_{j \in \mathcal{T}} n_{j,t} g_{ji} \tag{13}$$

In the economy, the number of firms grows at the rate η. At time t, the number of firms in the age cohort a is proportional to $e^{\eta(t-a)}$. As a result, the age distribution of firms is an exponential distribution. At the same time, as firms grow large, their growth rate converge. Therefore the size of a firm is a deterministic function of age, which gives rise to a Pareto distribution. The proposition below summarizes the details.

Proposition 4. *The right tail of the accumulative firm size distribution $F(n)$ is given by:*

$$F(n) = 1 - \alpha n^{-\frac{\eta}{\tau^*}}$$

where α is a constant.

Proof. See Appendix. □

The firm size distribution displays a heavy tail. The thickness of the right tail depends on the shape parameter, η/τ^*. A higher growth rate of the firm population, η, holding the firm growth rate τ^* constant, leads to smaller firm size differences across time and thus a thinner right tail. On the other hand, a higher firm growth rate, τ^*, given the population growth rate η fixed, results in a higher proportion of large firms and therefore a thicker right tail.

The firm size distribution found in this paper is consistent with the literature (Gabaix (2009)), but with a different economic mechanism. In this paper, firms grow over time at the same rate. The growing population of firms gives rise to different age cohorts, which, combined with firm growth, generate different size cohorts of firms. The relative size of the firm growth rate and the population growth rate determines the shape of the firm size distribution. A special case where $\tau^* = \eta$ will deliver Zipf's distribution. The fundamental linkage between economic growth and the firm size distribution makes the latter subject to the impacts of the same forces that influence the former.

4.3. Reducible Network

We have so far established an innovation model under an irreducible network and explored several implications of this network structure. In this section, the assumption that the technology network is irreducible will be relaxed. The key difference between an irreducible network and a reducible network is that, for the latter, technologies in one area need not communicate with technologies in all other areas. In terms of the network structure, this means that the network is more sparse than it was before. Knowledge spillovers in such a network display local effects instead of global effects. The definition below formally defines a reducible network.

Definition 2. A network is reducible if it can be partitioned exclusively into different subnetworks that are irreducible.

A reducible network represents an economy formed of multiple technology clusters. Each technology cluster is comprised of several sectors that are technically related to each other. Denote the i_{th} technology cluster by $\tilde{G}^i$ as follows:

$$\tilde{G}^i = \begin{bmatrix} \tilde{g}^i_{11} & \cdots & \tilde{g}^i_{1M_i} \\ \vdots & \ddots & \vdots \\ \tilde{g}^i_{Mi1} & \cdots & \tilde{g}^i_{M_iM_i} \end{bmatrix}$$

Every technology cluster is an irreducible subnetwork. Let the number of clusters be $\mathcal{N}$, then $\mathcal{N}M_i = M$, where M is the total number of sectors in the economy. The whole reducible network can be expressed as a matrix:

$$\tilde{G} = \begin{bmatrix} \tilde{G}^1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \tilde{G}^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \tilde{G}^{\mathcal{N}} \end{bmatrix}$$

where $\mathbf{0}$ is a zero matrix. Also define the sectoral innovation-adjusted cluster network as:

$$\tilde{\Phi}^i = \begin{bmatrix} \lambda^i_1\tilde{g}^i_{11} & \cdots & \lambda^i_1\tilde{g}^i_{1M_i} \\ \vdots & \ddots & \vdots \\ \lambda^i_{Mi}\tilde{g}^i_{Mi1} & \cdots & \lambda^i_{Mi}\tilde{g}^i_{M_iM_i} \end{bmatrix}$$

Within each technology cluster, most of the intuition developed previously in the case of an irreducible networks is retained. In particular, economic growth is determined by the subnetwork structure of each technology cluster as well as sectoral innovation rates. At the aggregate level, however, different clusters may experience different growth rates and economic growth is driven by the cluster that grows fastest. These findings are summarized by the following proposition.

Proposition 5. *If an economy is represented by a reducible technology network, $\tilde{G}$, then the long-run growth rates of each technology cluster and the economy are given by:*

$$\dot{N}^i_t/N^i_t = \tau^*_i$$

$$\dot{N}_t/N_t = \tau^{max} = \max_i\{\tau^*_i\}$$

where τ_i^* is the dominant eigenvalue of $\tilde{\Phi}^i$.

The reducible network structure allows sectors to grow at different rates across clusters but retain convergence within clusters. Note that it is possible for different clusters to grow at the same rate. Nonetheless, the observation that two clusters show the same growth rate does not necessarily imply that their underlying network structures are the same. It could be the case that one cluster has limited knowledge spillovers but higher sectoral innovation rates while the other cluster has strong knowledge spillovers but lower sectoral innovation rates. As an illustration, suppose that there are two technology clusters in the economy. These two technology clusters are represented by two 4×4 matrices as follows:

$$G^1 = \begin{bmatrix} 1 & 0.01 & 0.01 & 0.01 \\ 0.01 & 1 & 0.01 & 0.01 \\ 0.01 & 0.01 & 1 & 0.01 \\ 0.01 & 0.01 & 0.01 & 1 \end{bmatrix} G^2 = \begin{bmatrix} 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}$$

As a result, the network structure of the economy can be expressed as:

$$G = \begin{bmatrix} 1 & 0.01 & 0.01 & 0.01 & 0 & 0 & 0 & 0 \\ 0.01 & 1 & 0.01 & 0.01 & 0 & 0 & 0 & 0 \\ 0.01 & 0.01 & 1 & 0.01 & 0 & 0 & 0 & 0 \\ 0.01 & 0.01 & 0.01 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0.5 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 1 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}$$

Let $\lambda^1 = 0.1$ be the innovation rate for all sectors in cluster 1, and $\lambda^2 = 0.0412$ be the innovation rate for all sectors in cluster 2. It is easy to show that the dominant eigenvalues for both clusters are the same (0.1030), which means that the long-run growth rates for both clusters are the same.

In this example, sectors in cluster 1 are technologically isolated while those in cluster 2 are well connected. However, sectors in both clusters grow at the same rate. They grow due to different reasons. Sectors in cluster 1 grow mainly because they have a high innovation rate, while sectors in cluster 2 grow because they benefit from intersectoral knowledge spillovers. This also has implications for individual firms' growth dynamics. Firms initially rooted in some sector of cluster 1 are more likely to specialize in one area because the benefits of intersectoral knowledge spillovers are limited and firms have less incentive to internalize these benefits by expanding their portfolios. This is not the case for firms entering the second cluster. In this case, there is a higher probability that a firm will develop a more diversified portfolio.

The fact that clusters grow at different rates has natural implications for the dynamics of sector shares. The discrepancy of cluster growth rates leads to the shrinking of the shares of all clusters except the fastest growing cluster.[15] Therefore, there is no stationary sector shares. The shares of all sectors, other than those belong to the fastest growing cluster, vanish in the long run. However, the relative sector sizes within each cluster still approach constant in the long run.

To show the above results formally, we must first introduce some new notations. Define the vector of sector shares in cluster i at time t as $Q_t^i = (Q_{1,t}^i, Q_{2,t}^i, \cdots, Q_{Mi,t}^i)$. The aggregate share of a cluster at time t is thus $\widehat{Q}_t^i = \sum_h (Q_{h,t}^i)$, and the relative share of sector h within cluster i at time t is $\Pi_t(h|i) = \frac{Q_{h,t}^i}{\widehat{Q}_t^i}$. Denote the vector that contains the conditional sector shares within cluster i as $\Pi_t^i = (\Pi_t(1|i), \Pi_t(2|i), \cdots, \Pi_t(M_i|i))$. The next proposition shows that the conditional sector shares within cluster i converge to $\Pi^i = \lim_{t\to\infty} \Pi_t^i$ in the long run.

Proposition 6. *The conditional sector shares within cluster i, Π^i, satisfy the following system of equations:*

$$\psi_i^* \Pi^i = \Psi^i \Pi^i \ \ \forall i \in (1, 2, \cdots, \mathcal{N}),$$

where ψ_i^* is the dominant eigenvalue of Ψ^i. Except the fastest growing cluster, the shares of all other clusters shrink and eventually vanish:

$$\lim_{t\to\infty} \widehat{Q}_t^i \to 0, \ \ \forall i \neq i_{max}.$$

Comparing Proposition 6 with Proposition 3, the predictions of sector shares differ substantially under different assumptions of the network structure. When the economy is represented by a reducible network, the model predicts divergent sector shares, some of which become negligible. However, within each technology clusters, the conditional sector shares tend to remain non-negligible and stay constant over time. This is because sectors belong to the same technology cluster grow at the same rate in the long run. Therefore, although the absolute shares of sectors in the slow-growing clusters shrink over time compared to the absolute sector shares in the fast-growing clusters, the relative shares of sectors remain stable within each cluster in the long run.

The features of the reducible network also play an important role in shaping the distribution of firm sizes. To see this, note that in a reducible technology network, firms innovate locally within a cluster. They stochastically enter a sector and accumulate knowledge in that sector. They then apply their knowledge to other sectors where they have a comparative advantage. However, they will not be able to expand their portfolios in an unlimited fashion, because knowledge in one technology cluster may not be useful in others. In addition, the network structure of a cluster imposes some restrictions

[15] To simplify the analysis here, it is assumed that the fastest growing cluster is unique.

on how valuable each piece of knowledge is on average, thus providing a limit upon how fast a firm can innovate.

Given the impact of network structures on firm growth, the firm size distribution should be considered conditional on clusters. In particular, there are different firm size distributions over different clusters. Denote the accumulative firm size distribution in cluster i as $F_i(n)$. The following proposition characterizes the tail behavior of these distributions.

Proposition 7. *The right tail of the accumulative firm size distribution $F_i(n)$ for cluster i is given by:*

$$F_i(n) = 1 - \alpha_i n^{-\eta/\tau_i^*} \quad \forall i \in (1, 2, \cdots, \mathcal{N})$$

where τ_i^* is the dominant eigenvalue of $\tilde{\Phi}^i$ and α_i is a cluster specific constant.

As shown in Proposition 7, the cluster-specific growth rate determines the thickness of the right tail of the firm size distribution for each technology cluster. Proposition 7 provides a new angle to think about firm size differences compared to the traditional literature.[16] In the traditional literature, the main mechanism to generate a Pareto-type distribution is the assumption that firms follow a random growth process. In this case, firms in the same age cohort have different sizes due to different realizations of innovation shocks. Firms with good luck may expand and grow, while firms with bad luck shrink over time. In contrast, in our model, firms in the same age cohort can have different sizes if the underlying technologies that sustain their growth are different. Put differently, the fact that firms specialize in different technology clusters creates a fundamental difference in their growth potential, which, in turn, has an impact on their sizes.

To sum up, this section demonstrates that different network structures have important implications on the long term behaviors of the economy. In next section, a simple numerical analysis is conducted to highlight these differences. The purpose of such an exercise is to offer a visual comparison of the irreducible and reducible network structures.

4.4. Irreducible v.s. Reducible Network: A Numerical Analysis

Two network structures are constructed, an irreducible network and a reducible network. The irreducible network is represented by a 8×8 matrix, whose off-diagonal entries are generated randomly

[16]See, among others, Luttmer (2007, 2011).

from a uniform distribution with support (0,0.2), while the diagonal entries are normalized to 1.

$$G = \begin{bmatrix} 1 & 0.0844 & 0.1357 & 0.0554 & 0.0877 & 0.1419 & 0.1919 & 0.1782 \\ 0.1930 & 1 & 0.1515 & 0.0092 & 0.0763 & 0.1509 & 0.0681 & 0.1919 \\ 0.0315 & 0.1584 & 1 & 0.0194 & 0.1531 & 0.0552 & 0.1171 & 0.1094 \\ 0.1941 & 0.1919 & 0.0784 & 1 & 0.1590 & 0.1359 & 0.0448 & 0.0277 \\ 0.1914 & 0.1311 & 0.1311 & 0.1390 & 1 & 0.1310 & 0.1503 & 0.0299 \\ 0.0971 & 0.0071 & 0.0342 & 0.0634 & 0.0980 & 1 & 0.0510 & 0.0515 \\ 0.1601 & 0.1698 & 0.1412 & 0.1900 & 0.0891 & 0.0238 & 1 & 0.1681 \\ 0.0284 & 0.1868 & 0.0064 & 0.0069 & 0.1293 & 0.0997 & 0.1398 & 1 \end{bmatrix}$$

The sector-specific innovation rates are drawn from a uniform distribution with support (0,0.2), and the realization is:

$$\lambda = (0.1629, 0.1812, 0.0254, 0.1827, 0.1265, 0.0195, 0.0557, 0.1094)$$

An initial population of 4000 firms, 500 for each sector, are simulated, following the process specified by equation (9), for a total of 300 periods to generate the dynamics of firm development. In addition, a 4% growth rate of the population is assumed, and every entrant enters a sector following a constant probability distribution function $\tilde{p}$. $\tilde{p}$ is constructed to be equal to the equilibrium sector size distribution, which, in our example, is:

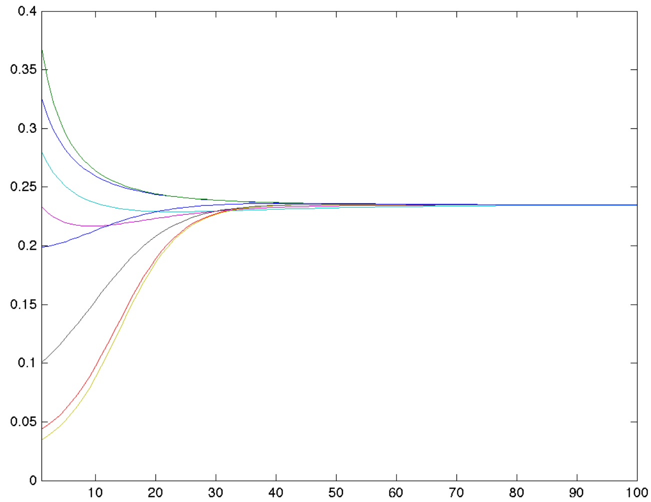
$$\tilde{p} = (0.2695, 0.3127, 0.0140, 0.1436, 0.1059, 0.0119, 0.0350, 0.1075)$$

The results below are not dependent on this assumption, which is simply employed to speed up the convergence. The simulation is applied to every generation of firms. The construction of an economy represented by a reducible network is similar. Assume that the economy now contains two 4×4 subnetworks in the diagonal blocks, each of which represents a technology cluster. The off-diagonal blocks are replaced with zero matrices. The intersectoral knowledge spillovers within each cluster of the reducible network are assumed to be the same as those in the irreducible network.

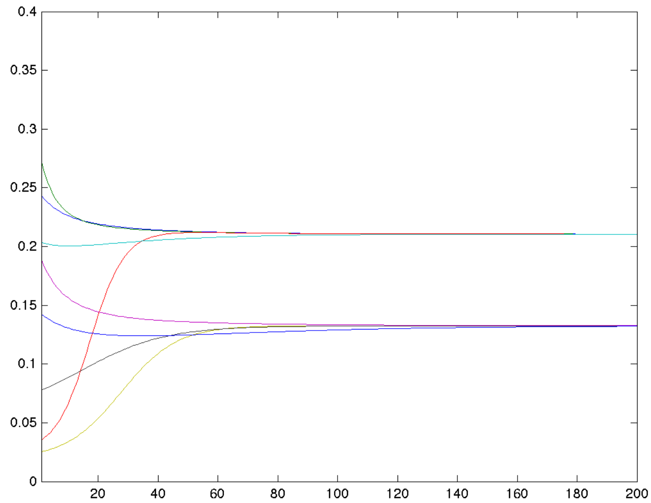
$$G' = \begin{bmatrix} 1 & 0.0844 & 0.1357 & 0.0554 & 0 & 0 & 0 & 0 \\ 0.1930 & 1 & 0.1515 & 0.0092 & 0 & 0 & 0 & 0 \\ 0.0315 & 0.1584 & 1 & 0.0194 & 0 & 0 & 0 & 0 \\ 0.1941 & 0.1919 & 0.0784 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0.1310 & 0.1503 & 0.0299 \\ 0 & 0 & 0 & 0 & 0.0980 & 1 & 0.0510 & 0.0515 \\ 0 & 0 & 0 & 0 & 0.0891 & 0.0238 & 1 & 0.1681 \\ 0 & 0 & 0 & 0 & 0.1293 & 0.0997 & 0.1398 & 1 \end{bmatrix}$$

To facilitate a fair comparison, all other parameters are equal to those used before. The simulation is done with the new network structure to generate a panel of firms for the reducible network.

Our first exercise here is to compare the long-run growth rate under these two different network structures. Using the simulated panel of firms, we are able to calculate the time series of sectoral growth rates. The results are shown in the figure below.



(a) IRREDUCIBLE



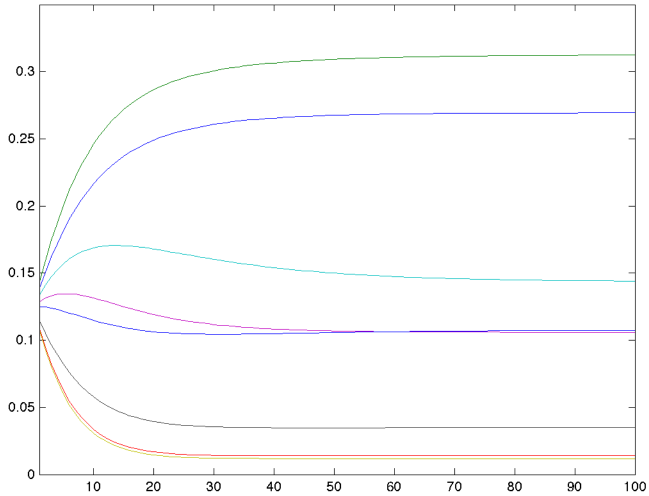
(b) REDUCIBLE

FIGURE 4

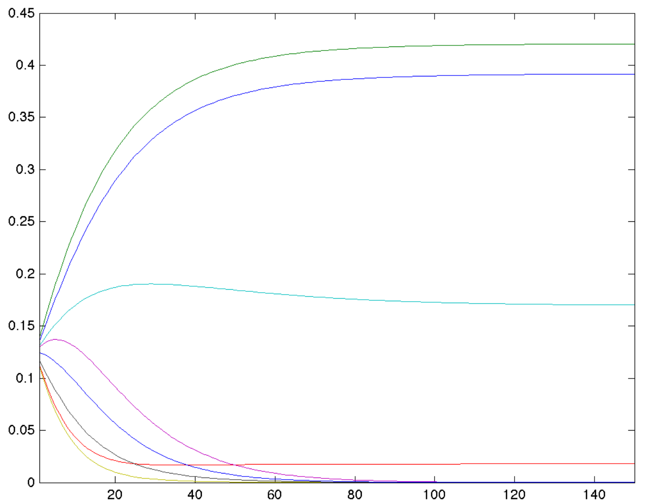
IRREDUCIBLE V.S. REDUCIBLE TECHNOLOGY NETWORK: SECTORAL GROWTH

The top panel of the Figure 4 depicts the dynamics of sectoral growth in the irreducible network. All sectors grow at the same rate in the long run, as predicted by Proposition 2. In contrast, sectors in the reducible network are divided into two clusters, as shown in the bottom panel of Figure 4. Sectors converge to the same growth rate in the long run within each cluster but diverge across clusters. As explained in Proposition 5, the divergence of sectoral growth rate in different clusters is driven by the facts that different technology clusters do not communicate with each other and that the network structures of different clusters are not the same.

Next, we analyze the change of sector shares over time in Figure 5. The comparison of the two panels in the figure reveals the significant difference of the sector share evolution. In the irreducible network, each sector's share share approaches a fixed number and stays constant in the long run.



(a) IRREDUCIBLE



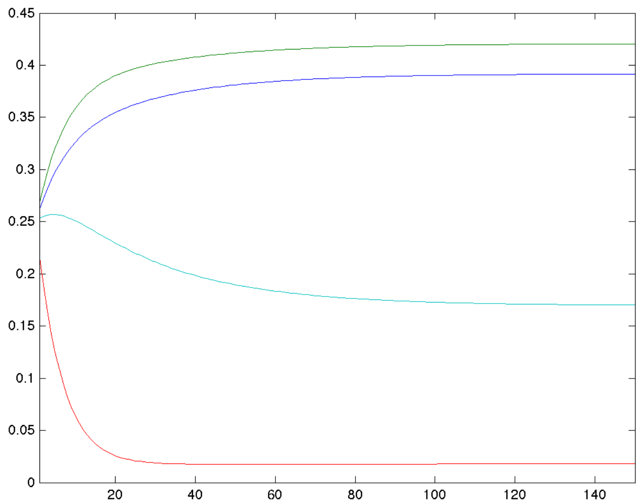
(b) REDUCIBLE

FIGURE 5

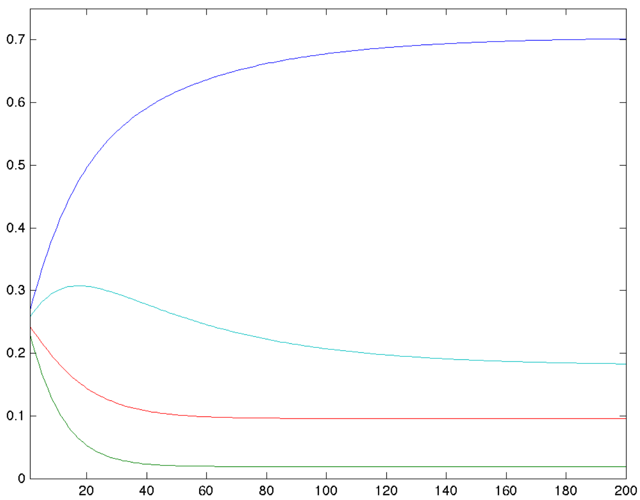
IRREDUCIBLE V.S. REDUCIBLE TECHNOLOGY NETWORK: SECTOR SHARE DISTRIBUTION

Nonetheless, in the reducible network, the shares of some sectors asymptotically approach zero, while others remain positive. The discrepancy of the dynamics of the sector shares for the two network structures is a direct result of heterogeneous sectoral growth. Sectors in the slow-growing cluster shrink over time relative to those in the fast-growing cluster.

Highlighted by Proposition 6, the relative sector shares within each technology cluster are constant in the reducible network in the long run. This feature is shown in FIGURE 6. Panel (a) and (b) in the figure corresponds to the top left and the bottom right blocks of the matrix G' respectively. In both panels of this figure, sectors start with the same share, and then change rapidly in their relative sizes, before converging to a stable state.



(a) CLUSTER 1

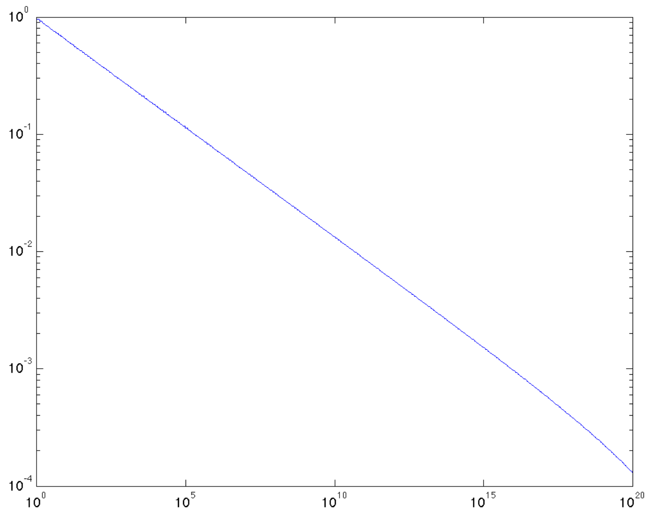


(b) CLUSTER 2

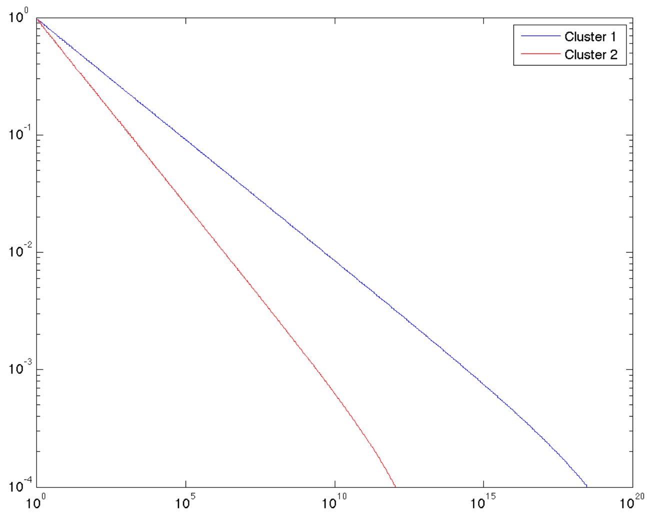
FIGURE 6

WITHIN CLUSTER SECTOR SHARE DISTRIBUTIONS

Lastly, we compare the firm size distributions in the irreducible and reducible networks in FIGURE 7. Note that the horizontal axis and the vertical axis are both in log scale. The firm size distributions in both cases are clearly log-linear, one notable feature of the Pareto distribution. For the reducible network, the firm size distribution in the two technology clusters differs in terms of the thickness of the right tails. As shown in Proposition 7, the thickness of the right tail is determined by the ratio of the population growth rate and the long-run growth rate of each technology cluster. For a given population growth rate, the technology cluster that grows faster displays a flatter firm size distribution, as is clearly the case in panel (b) of FIGURE 7.



(a) IRREDUCIBLE



(b) REDUCIBLE

FIGURE 7

IRREDUCIBLE V.S. REDUCIBLE TECHNOLOGY NETWORK: FIRM SIZE DISTRIBUTION

5. POLICY ANALYSIS

The theoretical framework established above reveals that different sectors contribute to economic growth differently. The higher a sector's innovation rate is, the more central the sector's position is in the technology network, the stronger its knowledge spillovers are and thus the more it contributes to economic growth. The aim of this section is to identify the importance of different sectors in terms of their ability to generate knowledge spillovers and stimulate economic growth. Towards this aim, we introduce a policy shock to a sector, holding everything else constant, so that the innovation rate of the sector increases by 1%. We then examine the impact of this policy on economic growth. Specifically, we consider two types of shocks: temporary and permanent policy shocks. The former lasts only for one period while the latter is permanent once implemented.

In order to proceed with the policy exercise, estimates for the parameters that determine the evolution of knowledge accumulation are needed. As a start point, we use the patent citation network as a proxy for the underlying technology network. The backward citation ratio is calculated to represent the strength of knowledge spillovers from the cited sector to the citing sector. We then proceed to calculate the sector-specific entry rate ηp_i. We first calculate the average annual growth rate of firm population from the NBER Patent Citation Data. The average growth rate is for the 11-year period between 1990 and 2000 where the data is most complete. The average probability of sector entry is derived for the same period. We exclude sector 33, Genetics, from the sample because this is a new sector that entered the dataset in 1977 and experienced volatile growth.[17]

The last set of parameters required are the sectoral innovation rates. They are constructed following equation (10). We calculate the increase in the number of patents per year due to the innovation of incumbents, and subtract that from the total annual increase of patent to get the number of patents from incumbents. Using the technology network, we infer the effective knowledge stocks for each sector which, combined with the previous results, gives rise to the sectoral innovation rate. The detailed parameter values are shown in TABLE A.4 in the Appendix.

With these parameters, we then simulate the dynamics of the economy, starting with 500 firms for each sector, for 800 periods. The temporary policy shock is introduced at period 400. The results are shown in FIGURE 8. The figure depicts the percentage increase in output for each sector shock relative to the status quo case. The horizontal axis represents the sector index. The figure shows significant heterogeneity across sectors. Policy shocks to most sectors have negligible impacts on aggregate output, with only a few exceptions.

[17] Including Genetics in the sample does not alter the results qualitatively.

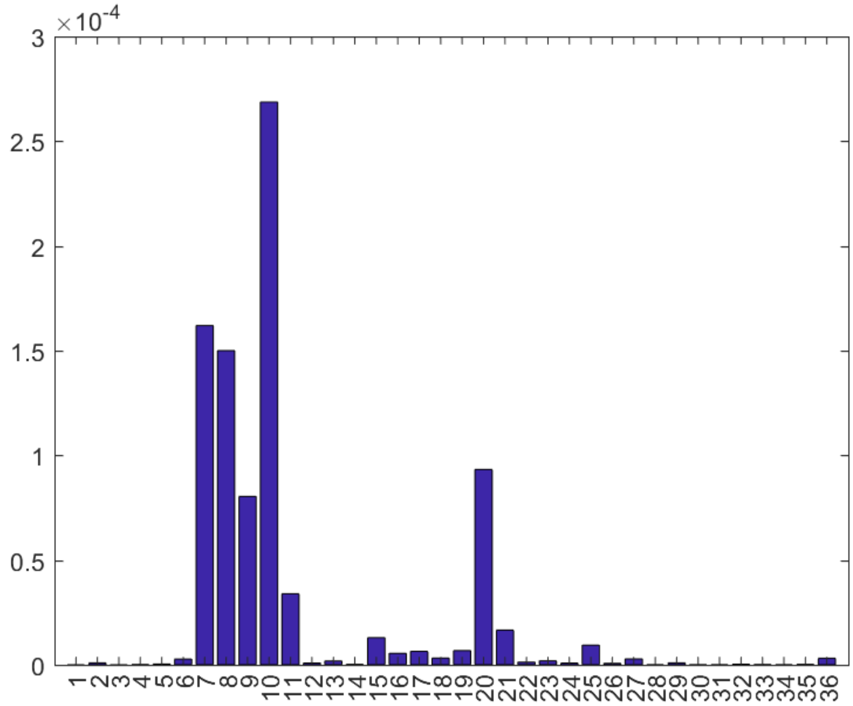


FIGURE 8

THE IMPACTS OF TEMPORARY POLICY SHOCKS AFFECTING EACH SECTOR

We list the top 4 and bottom 4 sectors in TABLE 3. The ranking is consistent with our expectation on which sector is important and which sector is not. For example, Information Storage is the most effective sector to target. A one-period 1% increase of the innovation rate in this sector leads to an increase in output of 0.27%. By contrast, sectors such as Gas and Agriculture are in the bottom of the ranking. It is no surprise that innovations in the Agriculture sector have little impact on other sectors since there are limited knowledge spillovers originating from this sector.

TABLE 3

SECTOR RESPONSES TO TEMPORARY POLICY SHOCKS: TOP 4 V.S. BOTTOM 4

Top 4 Sectors		Bottom 4 Sectors	
Sector Name	Output Increase (%)	Sector Name	Output Increase (%)
Information Storage	0.027	Apparel & Textile	0.0000089
Communication	0.016	Earth Working & Wells	0.0000073
Computer Hardware & Software	0.015	Gas	0.0000039
Semiconductor Devices	0.0093	Agriculture, Food & Textiles	0.0000023

We now turn to the case of permanent policy shocks. FIGURE 9 shows the effects of permanent policy shocks on economic growth in the long run. The vertical axis represents the growth rate change due to policy shocks. The values shown in the figure are calculated when the economy converges to a constant growth rate. Again, policy shocks to different sectors demonstrate considerable differences in their effects on economic growth. Depending on which sector is targeted, the impacts of a 1% increase of sectoral innovation rate on additional economic growth range from almost zero up to about 0.32%.

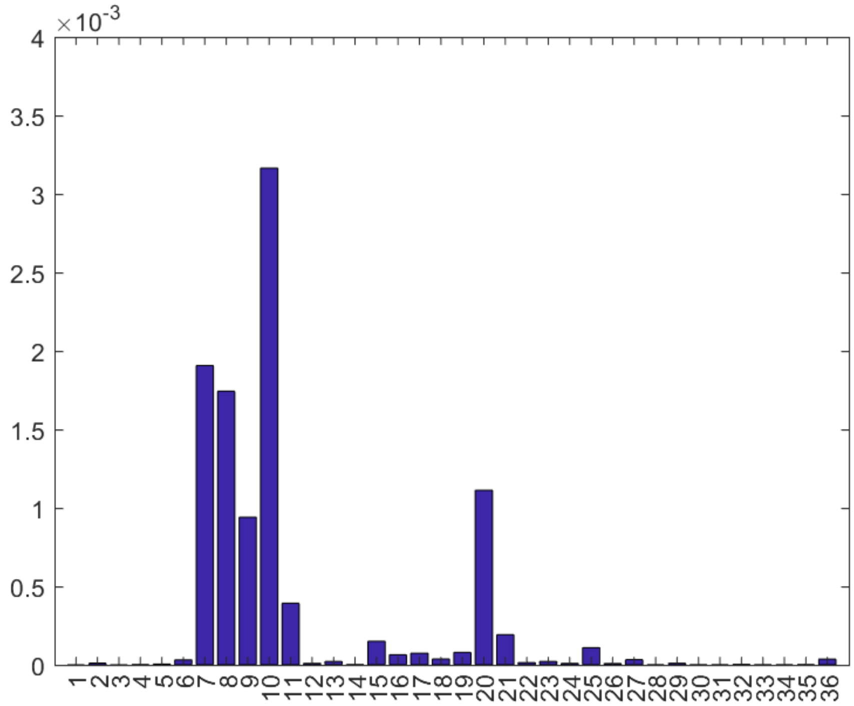


FIGURE 9

THE IMPACTS OF PERMANENT POLICY SHOCKS AFFECTING EACH SECTOR

TABLE 4 provides a counterpart of TABLE 3 for the permanent policy shocks. It is worthy to note that the ranking of the top 4 and the bottom 4 sectors here are the same as the ranking in TABLE 3. This result suggests that the importance of different sectors remains unchanged regardless of what type of policy is concerned.

TABLE 4

SECTOR RESPONSES TO PERMANENT POLICY SHOCKS: TOP 4 V.S. BOTTOM 4

Top 4 Sectors		Bottom 4 Sectors	
Sector Name	Growth Rate Change (%)	Sector Name	Growth Rate Change (%)
Information Storage	0.32	Apparel & Textile	0.00010
Communication	0.19	Earth Working & Wells	0.000085
Computer Hardware & Software	0.17	Gas	0.000045
Semiconductor Devices	0.11	Agriculture, Food & Textiles	0.000027

For the top 4 sectors in the table, permanent policy shocks have non-trivial effects on stimulating economic growth. For instance, the policy shock to Information Storage increases the economic growth rate by 0.32%, followed by 0.19% for Communication. As a comparison, the policy shock to the least responsive sector, Agriculture, Food & Textiles, contributes a 0.000027% increase of the economic growth rate. Thus, the most effective industry policy can generate an impact which is more than 10,000 times larger than the least effective industry policy, implying a strict hierarchy among sectors.

6. CONCLUSION

The paper provides novel empirical evidence on the importance of intersectoral knowledge spillovers for innovation using the patent citation data from U.S. Patent and Trademark Office. In particular, the paper documents that cross-sector knowledge spillovers are important for an individual firm's innovation in both the intensive margin as well as the extensive margin. In addition, the empirical evidence reveals that sectors differ in their efficiencies in utilizing cross-sector knowledge, reflecting significant sectoral heterogeneity.

Motivated by the empirical evidence documented in the paper, we construct a model of endogenous innovation on multiple technology sectors, where firms takes into account heterogeneous intersectoral knowledge spillovers when conducting innovation. The economy as a whole is modeled as a technology network that captures both intrasectoral and intersectoral knowledge spillovers. Firms enter the economy by producing a new patent in a particular sector and then accumulate knowledge to expand their patent portfolios to other sectors. The paper makes theoretical contributions to the understanding of the relationship between the technology network and a range of important economic issues, including economic growth, sector shares and the firm size distribution. Moreover, the paper demonstrates that a more sparse network structure limits the impacts of intersectoral knowledge spillovers, and shows how changes in the network structure affect the aggregate behavior of the economy.

The framework proposed in the paper offers a potential toolkit to identify key sectors in the economy. We evaluate a sector's importance by imposing an industry-oriented policy shock and simulate its impact on economic growth. The policy exercises show that there are enormous differences across sectors in regard to their contributions to economic growth. A marginal increase of the innovation rate for the most important sector can generate knowledge spillovers that result in additional economic growth which is 10,000 times larger than in the case of the least important sector.

Appendix

Proposition 1. The value function for the firm's problem has a solution as follows:

$$V(\tilde{\pi}^n) = \sum_{i\in\mathcal{T}}\sum_{s=1}^{n_i}\frac{\pi_{is}}{r} + \sum_{j\in\mathcal{T}} n_j R_j$$

where $(R_1 \cdots R_M)$ are the solutions to a system of M nonlinear equations:

$$rR_j = \sum_i \theta(\frac{\frac{\bar{\pi}_i}{r}+R_i}{\epsilon})^{\frac{1}{\epsilon-1}} g_{ji}[(1-\frac{1}{\epsilon})(\frac{\bar{\pi}_i}{r}+R_i)] \quad \forall j \in \mathcal{T}$$

And the optimal innovation choices are given by:

$$\lambda_i = \theta(\frac{\frac{\bar{\pi}_i}{r}+R_i}{\epsilon})^{\frac{1}{\epsilon-1}} \quad \forall i \in \mathcal{T}$$

Proof. Guess the value function $V(\tilde{\pi}_n) = \sum_{i\in\mathcal{T}}\sum_{s=1}^{n_i} a\pi_{is} + \sum_{j\in\mathcal{T}} n_j R_j$. Substitute the guessing value function form into the Bellman equation, then we have:

$$r(\sum_{i\in\mathcal{T}}\sum_{s=1}^{n_i} a\pi_{is} + \sum_{j\in\mathcal{T}} n_j R_j) = \max\Big\{\sum_{i\in\mathcal{T}}\sum_{s=1}^{n}\pi_{is} - \sum_{i\in\mathcal{T}} c(\lambda_i)\sum_j n_j g_{ji}$$
$$+\sum_{i\in\mathcal{T}}\lambda_i\sum_{j\in\mathcal{T}} n_j g_{ji}(a\bar{\pi}_i + R_i)\Big\}$$

The above equation holds if and only if:

$$a = \frac{1}{r} \tag{14}$$
$$r\sum_{j\in\mathcal{T}} n_j R_j = \max_{\{\lambda_i\}_i}\Big\{\sum_{i\in\mathcal{T}}\lambda_i\sum_{j\in\mathcal{T}} n_j g_{ji}(a\bar{\pi}_i + R_i) - \sum_{i\in\mathcal{T}} c(\lambda_i)\sum_j n_j g_{ji}\Big\} \tag{15}$$

First order condition with respect to λ_i gives:

$$c'(\lambda_i)\sum_j n_j g_{ji} = \sum_j n_j g_{ji}(a\bar{\pi}_i + R_i), \quad \forall i \in \mathcal{T}$$

which can be simplified to be:

$$c'(\lambda_i) = a\bar{\pi}_i + R_i$$

Combined with $a = 1/r$ and the cost function, we have the optimal innovation rate for each sector as follows::

$$\lambda_i^* = \theta(\frac{\bar{\pi}_i/r + R_i}{\epsilon})^{\frac{1}{\epsilon-1}}$$

Substitute the optimal innovation rates into (15)

$$r\sum_{j\in\mathcal{T}} n_j R_j = \sum_{i\in\mathcal{T}}\lambda_i^*\sum_{j\in\mathcal{T}} n_j g_{ji}(a\bar{\pi}_i + R_i) - \sum_{i\in\mathcal{T}} c(\lambda_i^*)\sum_j n_j g_{ji}$$

which holds if:

$$rn_jR_j = \sum_i \lambda_i^* n_j g_{ji}(a\bar{\pi}_i + R_i) - \sum_i c(\lambda_i^*) n_j g_{ji}$$

After some algebra, we have:

$$\begin{aligned}
rR_j &= \sum_i \lambda_i^* g_{ji}(\bar{\pi}_i/r + R_i) - \sum_i \lambda_i^{*\epsilon} \theta^{1-\epsilon} g_{ji} \\
&= \sum_i \theta(\frac{\bar{\pi}_i/r + R_i}{\epsilon})^{\frac{1}{\epsilon-1}} g_{ji}(\bar{\pi}_i/r + R_i) - \sum_i \theta^{\epsilon}(\frac{\bar{\pi}_i/r + R_i}{\epsilon})^{\frac{\epsilon}{\epsilon-1}} \theta^{1-\epsilon} g_{ji} \\
&= \theta \sum_i (\frac{\bar{\pi}_i/r + R_i}{\epsilon})^{\frac{1}{\epsilon-1}} g_{ji} \Big[\bar{\pi}_i/r + R_i - \frac{\bar{\pi}_i/r + R_i}{\epsilon}\Big] \\
&= \theta \sum_i (\frac{\bar{\pi}_i/r + R_i}{\epsilon})^{\frac{1}{\epsilon-1}} g_{ji} \Big[(1 - 1/\epsilon)(\frac{\bar{\pi}_i}{r} + R_i)\Big] \quad \forall j
\end{aligned}$$

This is what is required by the proposition.

□

Proposition 2. In the long run, different sectors converge to the same growth rate, which is equal to the aggregate growth rate of the economy:

$$\dot{N}_t/N_t = \dot{N}_{i,t}/N_{i,t} = \tau^* \text{ if } \eta < \tau^*$$
$$\dot{N}_t/N_t = \dot{N}_{i,t}/N_{i,t} = \eta \text{ if } \eta > \tau^*$$

where τ^* is the dominant eigenvalue of the matrix Φ.

Proof. Take the limit of $\Delta t \to 0$ of equation (10):

$$\begin{aligned}
\dot{N}_{i,t} &= L_t \eta p_i + \lambda_i \sum_{j \in \mathcal{T}} N_{j,t} g_{ji} \\
&= L_0 e^{\eta t} \eta p_i + \lambda_i \sum_{j \in \mathcal{T}} N_{j,t} g_{ji} \ \ \forall i
\end{aligned}$$

Normalize $L_0 = 1$ and the above equations in matrix form:

$$\begin{bmatrix} \dot{N}_{1,t} \\ \dot{N}_{2,t} \\ \vdots \\ \dot{N}_{M,t} \end{bmatrix} = \eta e^{\eta t} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix} + \begin{bmatrix} \lambda_1 g_{11} & \lambda_1 g_{21} & \cdots & \lambda_1 g_{M1} \\ \lambda_2 g_{12} & \lambda_2 g_{22} & \cdots & \lambda_2 g_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_M g_{1M} & \lambda_M g_{2M} & \cdots & \lambda_1 g_{MM} \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ \vdots \\ N_{M,t} \end{bmatrix}$$

Rewrite the equations in a compact form:

$$\dot{\widetilde{N}}_t = \eta e^{\eta t} \tilde{p} + \Phi \widetilde{N}_t$$

where

$$\dot{\widetilde{N}}_t = \begin{bmatrix} \dot{N}_{1,t} \\ \dot{N}_{2,t} \\ \vdots \\ \dot{N}_{M,t} \end{bmatrix}, \quad \tilde{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix}, \quad \widetilde{N}_t = \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ \vdots \\ N_{M,t} \end{bmatrix}$$

By the fundamental theorem of Picard and Lindelof, the above system admits a solution as follows:

$$\widetilde{N}_t = e^{\Phi t}\widetilde{N}_0 + \int_0^t e^{(t-s)\Phi}\eta e^{\eta s}\tilde{p}ds$$

Assume that Φ has distinct real eigenpairs (τ_1, V_1), $\cdots$, (τ_M, V_M), then $e^{\Phi t}$ can be decomposed to be as follows:

$$e^{\Phi t} = Ve^{Tt}V^{-1}$$

where

$$T = \begin{bmatrix} \tau_1 & 0 & \cdots & 0 \\ 0 & \tau_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \tau_M \end{bmatrix}, \quad V = aug(V_1, V_2, \cdots, V_M)$$

With the new notation, we have:

$$\widetilde{N}_t = Ve^{Tt}V^{-1}\widetilde{N}_0 + \int_0^t Ve^{(t-s)T}V^{-1}\eta e^{\eta s}\tilde{p}ds$$

Define $\Gamma = V^{-1}\widetilde{N}_0$, $H = \eta V^{-1}\tilde{p}$, we have for each sector:

$$\begin{aligned} N_{i,t} &= V_{i1}e^{\tau_1 t}\Gamma_1 + \cdots + V_{iM}e^{\tau_M t}\Gamma_M + \int_0^t e^{\eta s}[V_{i1}e^{\tau_1(t-s)}H_1 + \cdots + V_{iM}e^{\tau_M(t-s)}H_M]ds \\ &= V_{i1}e^{\tau_1 t}\Gamma_1 + \cdots + V_{iM}e^{\tau_M t}\Gamma_M + \sum_{j=1}^M \int_0^t V_{ij}e^{\tau_j t + (\eta - \tau_i)s}H_j ds \\ &= V_{i1}e^{\tau_1 t}\Gamma_1 + \cdots + V_{iM}e^{\tau_M t}\Gamma_M + \sum_{j=1}^M \frac{1}{\eta - \tau_j}V_{ij}e^{\tau_j t}[e^{(\eta-\tau_j)t} - 1]H_j \\ &= V_{i1}e^{\tau_1 t}\Gamma_1 + \cdots + V_{iM}e^{\tau_M t}\Gamma_M + \sum_{j=1}^M \frac{1}{\eta - \tau_j}V_{ij}(e^{\eta t} - e^{\tau_j t})H_j \end{aligned}$$

Define $\tau^* = \max_i\{\tau_i\}_i$. If $\eta > \tau^*$:

$$\begin{aligned} \lim_{t\to\infty} N_{i,t} &= \lim_{t\to\infty} e^{\eta t}\sum_j V_{ij}e^{(\tau_j-\eta)t}\Gamma_j + \sum_{j=1}^M \frac{1}{\eta-\tau_j}V_{ij}e^{\eta t}[1 - e^{(\tau_j-\eta)t}]H_j \\ &= e^{\eta t}\sum_{j=1}^M \frac{V_{ij}}{\eta - \tau_j}H_j \end{aligned}$$

From which we have:

$$\frac{\dot{N}_{i,t}}{N_{i,t}} = \eta, \quad \forall i$$

If $\eta < \tau^*$:

$$\lim_{t\to\infty} = \lim_{t\to\infty} e^{\tau^* t} \sum_{j=1}^{M} V_{ij} e^{(\tau_j - \tau^*)t} \Gamma_j + \sum_{j=1}^{M} \frac{1}{\eta - \tau_j} V_{ij} e^{\tau^* t} [e^{(\eta - \tau^*)t} - e^{(\tau_j - \tau^*)t}] H_j$$
$$= e^{\tau^* t} V_{ii^*} (\Gamma_{i^*} + \frac{V_{i^*}}{\tau^* - \eta} H_{i^*})$$

and

$$\frac{\dot{N}_{i,t}}{N_{i,t}} = \tau_{i^*}, \quad \forall i$$

Since all sectors grow at the same rate, so does the aggregate growth rate of the economy:

$$\frac{\dot{N}_t}{N_t} = \eta, \quad \text{if } \eta > \tau^*$$
$$\frac{\dot{N}_t}{N_t} = \tau^*, \quad \text{if } \eta < \tau^*$$

□

Proposition 3. The sector shares Q_t approach to a limit Q in the long run, which satisfy the following system of equations:

$$\psi^* Q = \Psi Q$$

where $\psi^* = \tau^* + 1$.

Proof.

$$\widetilde{N}_{t+1} = L_t \eta \tilde{p} + \Psi \widetilde{N}_t$$

Divide both sides of the equation by N_t:

$$\frac{1}{N_t} \widetilde{N}_{t+1} = \frac{L_t \eta}{N_t} \tilde{p} + \Psi Q_t$$

Which can be rewritten as:

$$\frac{1}{N_{t+1}} \frac{N_{t+1}}{N_t} \widetilde{N}_{t+1} = \frac{L_t \eta}{N_t} \tilde{p} + \Psi Q_t$$

Note that:

$$\frac{N_{t+1}}{N_t} = 1 + \frac{L_t \eta + \sum_i \lambda_i \sum_j N_{j,t} g_{ji}}{N_t}$$
$$= 1 + \frac{L_t \eta}{N_t} + \frac{\sum_i \lambda_i \sum_j N_{j,t} g_{ji}}{N_t}$$

As shown in Proposition 2, as $t \to \infty$, the innovation from the existing knowledge stocks dominates the innovation from the entrants so we have:

$$\lim_{t\to\infty} \frac{\sum_i \lambda_i \sum_j N_{j,t} g_{ji}}{N_t} = \tau^*$$

Plug the above result to the previous equation:

$$\frac{N_{t+1}}{N_t} = 1 + \frac{L_t\eta}{N_t} + \tau^*$$

With the result and note the fact that $\frac{1}{N_{t+1}}\widetilde{N}_{t+1} = Q_{t+1}$, we have:

$$(1 + \frac{L_t\eta}{N_t} + \tau^*)Q_{t+1} = \frac{L_t\eta}{N_t}\tilde{p} + \Psi Q_t$$

Rearrange the above equation to get:

$$(1 + \tau^*)Q_{t+1} = \frac{L_t\eta}{N_t}(\tilde{p} - Q_{t+1}) + \Psi Q_t$$

Take the limit of $t \to \infty$ for the above equation:

$$\lim_{t\to\infty}(1 + \tau^*)Q_{t+1} = \lim_{t\to\infty}\frac{L_t\eta}{N_t}(\tilde{p} - Q_{t+1}) + \lim_{t\to\infty}\Psi Q_t$$

Note that $\lim_{t\to\infty}\frac{L_t\eta}{N_t} = 0$ when $\tau^* > \eta$, so we have:

$$\psi^* Q = \Psi Q$$

Where $\psi^* = 1 + \tau^*$. □

Proposition 4. The right tail of the accumulative firm size distribution $F(n)$ is given by:

$$F(n) = 1 - \alpha n^{-\frac{\eta}{\tau^*}}$$

where α is a constant.

Proof. Conditional on a firm entering sector i at the beginning, start from equation (13) and take the limit of $\Delta t \to 0$,

$$\dot{n}_t = \sum_i \lambda_i \sum_j n_{j,t} g_{ji}$$

where $n_{i,0} = 1$ and $n_{j,0} = 0\ \forall j \neq i$. We already know from Proposition 2 that the aggregate number of patents N_t grows a rate τ^* in the long run. This will be true for individual firms as well because of the constant return to scale of production function.

$$\lim_{t\to\infty}\dot{n}_t = \tau^* n_t$$

The above equation implies that firm size for the same cohort of firms that enter the same sector at the beginning is a deterministic function of age:

$$n_t = \alpha_i e^{\tau^* t}$$

where α_i is a constant specific to initial sector that firms enter. Revert the above equation to express a firm's age in terms of size:

$$t = \frac{1}{\tau^*} \log \frac{n_t}{\alpha_i}$$

Recall the population of firms grows at the rate η, so the distribution of firm age is an exponential distribution. The proportion of firms older than a is thus:

$$Prob(\text{firms older than a}) = e^{-\eta a}$$

Substitute firm age a in terms of firm size n, we have:

$$Prob(\text{firms larger than n}) = e^{-\frac{\eta}{\tau^*}\log\frac{n}{\alpha_i}} = (\frac{n}{\alpha_i})^{-\frac{\eta}{\tau^*}}$$

The accumulative firm size distribution, conditional on firms entering sector i at the beginning, is thus:

$$F_i(n) = 1 - (\frac{n}{\alpha_i})^{-\frac{\eta}{\tau^*}}$$

The unconditional firm size distribution is therefore:

$$\begin{aligned} F(n) &= \sum_i p_i F_i(n) \\ &= \sum_i p_i(1 - (\frac{n}{\alpha_i})^{-\frac{\eta}{\tau^*}}) \\ &= 1 - \sum_i p_i(\frac{1}{\alpha_i})^{-\frac{\eta}{\tau^*}} n^{-\frac{\eta}{\tau^*}} \\ &= 1 - \alpha n^{-\frac{\eta}{\tau^*}} \end{aligned}$$

where $\alpha = \sum_i p_i(\frac{1}{\alpha_i})^{-\frac{\eta}{\tau^*}}$ □

Proposition 5. If an economy is represented by a reducible technology network, $\tilde{G}$, then the long-run growth rates of each technology cluster and the economy are given by:

$$\dot{N}_t^i / N_t^i = \tau_i^*$$

$$\dot{N}_t / N_t = \tau^{max} = \max_i \{\tau_i^*\}$$

where τ_i^* is the dominant eigenvalue of $\tilde{\Phi}^i$.

Proof. Within each cluster, the proof is the same as the proof of Proposition 2. Since each cluster is a irreducible network, all sectors within a cluster grow at the same rate, which is determined by the eigenvalue τ_i^* of the network of every cluster Φ_i.

Since different clusters grow at different rate, as time passes on, the fastest growing cluster will dominate and the aggregate growth of the economy is thus determined by the growth of the dominant cluster. □

Proposition 6. The conditional sector shares within cluster i, Π^i, satisfy the following system of equations:

$$\psi_i^* \Pi^i = \Psi^i \Pi^i \ \ \forall i \in (1, 2, \cdots, \mathcal{N}),$$

where ψ_i^* is the dominant eigenvalue of Ψ^i. Except the fastest growing cluster, the shares of all other clusters shrink and eventually vanish:

$$\lim_{t\to\infty} \widehat{Q}_t^i \to 0, \ \ \forall i \neq i_{max}.$$

Proof. From results in Proposition 5, we know that:

$$N_t^i = c_i e^{\tau_i^* t} \ \ \forall i \in (1, \cdots, \mathcal{N})$$
$$N_t = c e^{\tau^{max} t}$$

where c_i and c are constants specific to cluster i and the whole economy respectively. Since $\tau_i^* < \tau_{max}$, we have:

$$\lim_{t\to\infty} \frac{N_t^i}{N_t} = \lim_{t\to\infty} \frac{c_i}{c} e^{(\tau_i^* - \tau_{max})t} = 0$$

so

$$\lim_{t\to\infty} \widehat{Q}_t^i = \lim_{t\to\infty} \frac{\sum_h N_{h,t}^i}{N_t} = \lim_{t\to\infty} \frac{N_t^i}{N_t} = 0 \ \ \forall i \neq i_{max}$$

It is easy to see from the above results that sectors other than those in the fastest growing cluster shrink in terms of the relative size. However, we can still explore the relative size distribution of sectors within each cluster. Due to the fact that each cluster consists of an irreducible network, we can examine the within cluster sector size distribution separately for each cluster. Specifically, for cluster i:

$$\widetilde{N}_{t+1}^i = L_t \eta \tilde{p}^i + \Psi^i \widetilde{N}_t^i$$

Following the similar procedure in the proof of Proposition 3, we have:

$$(1 + \frac{L_t \eta \sum_{h=1}^{M_i} p_h^i}{N_t^i} + \tau_i^*)\Pi_{t+1}^i = \frac{L_t \eta}{N_t} \tilde{p}^i + \Psi^i \Pi_t^i$$

Note that:

$$\lim_{t\to\infty} \frac{L_t \eta}{N_t} = 0$$

Therefore, when $t \to \infty$, the above system of equations simplify to:

$$\psi_i^* \Pi^i = \Psi^i \Pi^i$$

where $\psi^* = 1 + \tau^*$. □

Proposition 7. The right tail of the accumulative firm size distribution $F_i(n)$ for cluster i is given by:

$$F_i(n) = 1 - \alpha_i n^{-\eta/\tau_i^*} \quad \forall i \in (1, 2, \cdots, \mathcal{N})$$

where τ_i^* is the dominant eigenvalue of $\tilde{\Phi}^i$ and α_i is a cluster specific constant.

Proof. The proof here follows that in Proposition 4. The population growth rate of firms is the same across clusters, so we have the same firm age distribution over different clusters. The only difference is that we have cluster specific growth rate, so firms in the same age cohort can have different sizes depending on which cluster they belong to. This feature determines the distribution of firm sizes in different clusters. In particular, $\frac{\eta}{\tau_i^*}$ is the shape parameter for the firm size distribution of each cluster, which give rises to our results in the proposition. □

TABLE A.1

SUMMARY OF TECHNOLOGY CATEGORIES

Category by HJB	Relabel	Name	Number
11	1	Agriculture,Food,Textiles	7808
12	2	Coating	21257
13	3	Gas	6733
14	4	Organic Compounds	49041
15	5	Resins	53567
19	6	Miscellaneous-Chemical	141561
21	7	Communications	101276
22	8	Computer Hardware & Software	86433
23	9	Computer Peripherials	30084
24	10	Information Storage	44174
25	11	Electronic business methods and software	13882
31	12	Drugs	97507
32	13	Surgery & Med Inst.	53455
33	14	Genetics	4480
39	15	Miscellaneous-Drug& Med	9035
41	16	Electrical Devices	52782
42	17	Electrical Lighting	23333
43	18	Measuring & Testing	46449
44	19	Nuclear & X-rays	22562
45	20	Power Systems	51666
46	21	Semiconductor Devices	55702
49	22	Miscellaneous-Elec	32256
51	23	Mat. Proc & Handling	56574
52	24	Metal Working	35482
53	25	Motors & Engines + Parts	36162
54	26	Optics	16837
55	27	Transportation	30674
59	28	Miscellaneous-Mechanical	49743
61	29	Agriculture,Husbandry,Food	21358
62	30	Amusement Devices	8817
63	31	Apparel & Textile	11825
64	32	Earth Working & Wells	21464
65	33	Furniture,House Fixtures	18616
66	34	Heating	13527
67	35	Pipes & Joints	10707
68	36	Receptacles	20888
69	37	Miscellaneous-Others	98724

The first column corresponds to the original numeric technology category defined by Hall, Jaffe and Trajjtenberg(HJT). The second column is the relabelled category number in the paper and the last column shows the number of patents in each technology category.

TABLE A.2

PROBIT REGRESSIONS ACROSS SECTORS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Mproximity^{x}_{f,t-1}$	49.21***	22.26***	37.60***	17.99***	17.49***	14.97***	14.82***	8.05***
	9.33	2.1	7.93	1.11	1.04	0.51	0.71	0.4
$y^{x}_{f,t-1}$	1.78***	1.23***	1.51***	1.74***	1.61***	1.52***	1.62***	1.30***
	0.1	0.04	0.08	0.05	0.04	0.03	0.04	0.03
$N_{f,t-1}$	0.01*	0.03***	0.01	0	0.01*	-0.02***	0.01	0.01
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-3.05***	-2.61***	-2.96***	-2.66***	-2.70***	-1.86***	-2.37***	-2.49***
	0.07	0.05	0.07	0.05	0.05	0.04	0.04	0.04
Observations	55,194	55,194	55,194	55,194	55,194	55,194	55,194	55,194

TABLE A.2 Continued

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
$Mproximity^{x}_{f,t-1}$	23.54***	18.07***	20.08***	8.75***	12.86***	233.11***	32.34***	23.42***
	1.78	1.26	1.76	0.45	0.97	24.73	2.92	1.63
$y^{x}_{f,t-1}$	1.49***	1.71***	1.28***	2.30***	2.12***	2.04***	1.96***	1.48***
	0.06	0.06	0.06	0.05	0.05	0.1	0.08	0.04
$N_{f,t-1}$	0.01	0	0.02**	-0.01*	-0.01	0.03***	0.01	0.01
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-3.00***	-2.87***	-3.29***	-2.48***	-2.23***	-3.50***	-2.82***	-2.49***
	0.06	0.06	0.08	0.05	0.05	0.12	0.07	0.05
Observations	55,194	55,194	55,194	55,194	55,194	55,194	55,194	55,194

TABLE A.2 Continued

	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
$Mproximity^{x}_{f,t-1}$	39.98***	14.24***	24.61***	20.92***	14.13***	47.67***	26.08***	43.16***
	4.66	1.28	2.82	1.38	1.33	3.07	1.18	2.42
$y^{x}_{f,t-1}$	1.57***	1.20***	1.35***	1.44***	1.74***	1.36***	1.41***	1.36***
	0.05	0.03	0.05	0.04	0.06	0.05	0.03	0.04
$N_{f,t-1}$	0.03***	0.03***	0.03***	0.02***	0.03***	0.02***	0.01*	0.02***
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.69***	-2.18***	-2.47***	-2.48***	-2.93***	-2.52***	-1.99***	-2.26***
	0.06	0.04	0.05	0.05	0.06	0.05	0.04	0.04
Observations	55,194	55,194	55,194	55,194	55,194	55,194	55,194	55,194

TABLE A.2 Continued

	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
$Mproximity^{x}_{f,t-1}$	27.63***	23.76***	49.76***	34.52***	43.78***	182.38***	75.24***	21.80***
	1.88	2.66	3.07	1.52	3.25	17.95	7.24	2.52
$y^{x}_{f,t-1}$	1.63***	1.52***	1.95***	1.52***	2.07***	2.32***	1.84***	2.15***
	0.05	0.06	0.05	0.04	0.06	0.13	0.07	0.07
$N_{f,t-1}$	0.01	0.04***	0	-0.01	-0.01**	0.01	0.02***	-0.01
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.50***	-2.88***	-2.63***	-1.98***	-2.54***	-2.97***	-2.72***	-2.64***
	0.05	0.07	0.06	0.04	0.06	0.1	0.06	0.06
Observations	55,194	55,194	55,194	55,194	55,194	55,194	55,194	55,194

TABLE A.2 Continued

	(33)	(34)	(35)	(36)	(37)
$Mproximity^x_{f,t-1}$	41.57***	80.25***	93.52***	44.97***	15.37***
	2.78	8.48	9.01	2.89	0.52
$y^x_{f,t-1}$	1.76***	1.60***	1.37***	1.62***	1.40***
	0.05	0.07	0.06	0.05	0.03
$N_{f,t-1}$	-0.01	0.02**	0.02***	-0.01*	-0.01**
	0.01	0.01	0.01	0.01	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.47***	-2.73***	-2.87***	-2.30***	-1.70***
	0.05	0.06	0.06	0.05	0.03
Observations	55,194	55,194	55,194	55,194	55,194

The results in this table are based on the estimation of Probit regression (2) for each sector. Note that there is no sector dummies here since the regression itself is sector-specific.

TABLE A.3

NEGATIVE BINOMIAL REGRESSIONS ACROSS SECTORS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Wpatent^{x}_{f,t-1}$	0.58***	0.42***	0.46***	0.65***	0.55***	0.24***	0.40***	0.66***
	0.05	0.03	0.05	0.03	0.03	0.02	0.03	0.03
$I(x_{f,t-1} = 1)$	2.89***	1.86***	2.33***	1.77***	2.15***	1.70***	2.34***	1.70***
	0.2	0.1	0.16	0.1	0.08	0.05	0.08	0.07
$N_{f,t-1}$	-0.01	0.02**	0.01	-0.02***	-0.02***	0.02***	-0.02**	-0.02**
	0.02	0.01	0.01	0	0.01	0	0.01	0.01
$PatentCounts^{x}_{f,t-1}$	0.11***	0.07***	0.07	0.07***	0.10***	0.11***	0.09***	0.08***
	0.04	0.02	0.06	0.01	0.01	0.01	0.01	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.32***	-2.26***	-2.84***	-1.36***	-1.79***	-1.62***	-2.00***	-1.87***
	0.27	0.12	0.23	0.09	0.09	0.05	0.1	0.08
$ln\alpha$	0.81***	0.19**	0.88***	-0.26***	0.16**	-0.15***	0.36***	0.26***
	0.26	0.09	0.15	0.1	0.07	0.05	0.07	0.07
Observations	28,437	28,385	28,435	28,294	28,253	27,898	28,041	28,134

TABLE A.3 Continued

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
$Wpatent^x_{f,t-1}$	0.59***	0.60***	0.72***	0.32***	0.12***	0.63***	0.61***	0.42***
	0.06	0.06	0.07	0.02	0.03	0.03	0.04	0.04
$I(x_{f,t-1} = 1)$	2.59***	2.87***	1.86***	3.18***	2.92***	2.14***	2.31***	2.28***
	0.15	0.21	0.21	0.09	0.12	0.22	0.19	0.08
$N_{f,t-1}$	-0.02*	-0.04***	-0.02*	-0.04***	0.02**	-0.02*	-0.04***	-0.03***
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$PatentCounts^x_{f,t-1}$	0.07***	0.07***	0.07**	0.09***	0.09***	0.17**	0.18***	0.11***
	0.02	0.01	0.03	0.01	0.01	0.07	0.04	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.54***	-2.37***	-2.60***	-2.35***	-3.02***	-2.17***	-2.63***	-1.89***
	0.18	0.21	0.32	0.11	0.11	0.27	0.2	0.13
$ln\alpha$	0.73***	0.70***	0.57**	-0.23***	0.14	0.65***	0.84***	0.33***
	-0.12	0.13	0.27	0.07	0.11	0.2	0.14	0.09
Observations	28,282	28,240	28,383	27,820	28,058	28,195	28,399	28,225

TABLE A.3 Continued

	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
$Wpatent^x_{f,t-1}$	0.30***	0.26***	0.33***	0.45***	0.39***	0.32***	0.19***	0.26***
	0.05	0.03	0.05	0.03	0.03	0.04	0.02	0.03
$I(x_{f,t-1} = 1)$	2.57***	1.89***	2.10***	1.93***	2.89***	1.93***	1.75***	1.87***
	0.11	0.07	0.11	0.07	0.14	0.1	0.06	0.08
$N_{f,t-1}$	0.03**	0.03***	0.04***	0	0.01	0.03***	0.02***	0.03***
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$PatentCounts^x_{f,t-1}$	0.04**	0.07***	0.06***	0.10***	0.09***	0.10***	0.12***	0.07***
	0.02	0.01	0.02	0.01	0.02	0.02	0.01	0.02
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.69***	-2.09***	-2.39***	-1.99***	-3.02***	-2.44***	-2.10***	-2.19***
	0.24	0.1	0.18	0.1	0.14	0.14	0.08	0.11
$ln\alpha$	0.70***	0.13*	0.46***	0.32***	0.51***	0.59***	0.12*	0.25***
	0.14	0.08	0.14	0.07	0.12	0.1	0.06	0.09
Observations	28,369	28,286	28,371	28,254	28,226	28,325	28,224	28,331

TABLE A.3 Continued

	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
$Wpatent^x_{f,t-1}$	0.36***	0.43***	0.18***	0.22***	0.19***	0.09	0.21***	0.25***
	0.03	0.04	0.03	0.03	0.03	0.09	0.05	0.04
$I(x_{f,t-1} = 1)$	2.42***	2.40***	2.88***	1.83***	3.03***	4.49***	2.63***	3.65***
	0.09	0.14	0.12	0.07	0.13	0.54	0.2	0.15
$N_{f,t-1}$	-0.01	0.03**	0.02**	0	0.01	0.08***	0.04***	0.01
	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01
$PatentCounts^x_{f,t-1}$	0.10***	0.08***	0.08***	0.11***	0.10***	0.04	0.13**	0.11***
	0.02	0.02	0.02	0.01	0.03	0.13	0.06	0.02
$YearDummy$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.21***	-2.75***	-3.24***	-2.00***	-3.04***	-5.70***	-3.43***	-3.45***
	0.13	0.17	0.15	0.09	0.16	0.46	0.25	0.18
$ln\alpha$	0.41***	0.48***	0.66***	0.34***	0.53***	1.32***	1.08***	0.32*
	0.1	0.16	0.14	0.07	0.18	0.43	0.16	0.19
Observations	28,318	28,393	28,300	28,246	28,337	28,412	28,382	28,352

TABLE A.3 Continued

	(33)	(34)	(35)	(36)	(37)
$Wpatent^x_{f,t-1}$	0.26***	0.21***	0.36***	0.32***	0.14***
	0.03	0.04	0.04	0.03	0.02
$I(x_{f,t-1} = 1)$	2.27***	2.51***	1.88***	2.06***	1.53***
	0.13	0.14	0.15	0.1	0.05
$N_{f,t-1}$	0.01	0.04***	0.01	-0.01	0.02***
	0.01	0.01	0.01	0.01	0.01
$PatentCounts^x_{f,t-1}$	0.07**	0.05	0.09*	0.10***	0.11***
	0.03	0.04	0.05	0.02	0.01
$YearDummy$	Yes	Yes	Yes	Yes	Yes
$Firm\text{-}sector\ FE$	Yes	Yes	Yes	Yes	Yes
$Constant$	-2.79***	-3.19***	-2.70***	-2.15***	-1.76***
	0.16	0.22	0.2	0.13	0.05
$ln\alpha$	0.82***	0.71***	0.69***	0.49***	-0.18***
	0.09	0.15	0.13	0.1	0.05
Observations	28,386	28,422	28,427	28,387	28,134

The results in this table are based on the estimation of Negative Binomial regression (3) for each sector. Note that there is no sector dummies here since the regression itself is sector-specific.

TABLE A.4

SECTORAL ENTRY PROBABILITY AND INNOVATION INTENSITY

Name	Entry Probability	Innovation Intensity
Agriculture,Food,Textiles	0.001078	0.056423
Coating	0.003607	0.069459
Gas	0.001191	0.044832
Organic Compounds	0.002537	0.076724
Resins	0.003671	0.069619
Miscellaneous-Chemical	0.020506	0.062229
Communications	0.012182	0.105481
Computer Hardware & Software	0.010279	0.095151
Computer Peripherials	0.002868	0.108225
Information Storage	0.003202	0.121202
Electronic business methods and software	0.004119	0.097338
Drugs	0.014888	0.088518
Surgery & Med Inst.	0.011166	0.076118
Miscellaneous-Drug& Med	0.00267	0.65657
Electrical Devices	0.004401	0.08409
Electrical Lighting	0.003373	0.083668
Measuring & Testing	0.006807	0.072112
Nuclear & X-rays	0.002453	0.078219
Power Systems	0.005826	0.074447
Semiconductor Devices	0.00184	0.10364
Miscellaneous-Elec	0.005005	0.086706
Mat. Proc & Handling	0.013088	0.058572
Metal Working	0.006256	0.067428
Motors & Engines + Parts	0.005509	0.063773
Optics	0.002104	0.081174
Transportation	0.006851	0.062832
Miscellaneous-Mechanical	0.011863	0.065531
Agriculture,Husbandry,Food	0.006891	0.054318
Amusement Devices	0.00407	0.072859
Apparel & Textile	0.00418	0.057728
Earth Working & Wells	0.004055	0.039236
Furniture,House Fixtures	0.006618	0.063958
Heating	0.003005	0.048228
Pipes & Joints	0.00248	0.061579
Receptacles	0.006259	0.06207
Miscellaneous-Others	0.026874	0.066039

References

Acemoglu, D., Akcigit, U., Bloom, N., and Kerr, W. R. (2013). Innovation, reallocation and growth. *NBER working paper*.

Acemoglu, D., Akcigit, U., and Kerr, W. R. (2016). Innovation network. *Proceedings of the National Academy of Sciences*, 113(41):11483–11488.

Aghion, P., Harris, C., and Vickers, J. (1997). Competition and growth with step-by-step innovation: An example. *European Economic Review*, 41(3-5):771 – 782.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, (2):323–351.

Akcigit, U. and Kerr, W. R. (forthcoming). Growth through heterogeneous innovations. *Journal of Political Economy*.

Amir, R. and Wooders, J. (1999). Effects of one-way spillovers on market shares, industry price, welfare, and R&D cooperation. *Journal of Economics & Management Strategy*, 8(2):223–249.

Anbarci, N., Lemke, R., and Roy, S. (2002). Inter-firm complementarities in R&D: a re-examination of the relative performance of joint ventures. *International Journal of Industrial Organization*, 20(2):191–213.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, (4):1347–1393.

Blundell, R., Griffith, R., and van Reenen, J. (1999). Market share, market value and innovation in a panel of british manufacturing firms. *Review of Economic Studies*, 66(3):529 – 554.

Cai, J. and Li, N. (2012). Growth through intersectoral knowledge linkages. In *Mimeo: International Monetary Fund*.

D'Aspremont, C. and Jacquemin, A. (1988). Cooperative and noncooperative R&D in duopoly with spillovers. *The American Economic Review*, 78(5):1133–1137.

Erkal, N. and Piccinin, D. (2010). Cooperative R&D under uncertainty with free entry. *International Journal of Industrial Organization*, 28(1):74 – 85.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *The Quarterly Journal of Economics*, (3):739.

Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics*, pages 255–293.

Grossman, G. M. and Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies*, 58(1):43 – 61.

Hall, B. H., Jaffe, A., and Trajtenberg, M. (2005). Market value and patent citations. *The RAND Journal of Economics*, (1):16–38.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Jackson, M. O. and Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *The American Economic Review*, (3):890.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *The American Economic Review*, (5):984–1001.

Klette, T. J. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986 – 1018.

Kortum, S. S. (1997). Research, patenting, and technological change. *Econometrica*, (6):1389–1419.

Lentz, R. and Mortensen, D. T. (2005). Productivity growth and worker reallocation. *International Economic Review*, (3):731–749.

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

Luttmer, E. G. J. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, (3):1103–1144.

Luttmer, E. G. J. (2011). On the mechanics of firm growth. *Review of Economic Studies*, 78(3):1042 – 1068.

Oberfield, E. (2012). Business networks, production chains and productivity: A theory of input-output architecture.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, (5):1002–1037.

Suzumura, K. (1992). Cooperative and noncooperative R& D in an oligopoly with spillovers. *The American Economic Review*, 82(5):1307–1320.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA : MIT Press, c2010.