TECHNOLOGY NETWORK INNOVATION AND DISTRIBUTION

Jingong (Jim) Huang *

University of Melbourne

Motivated by empirical evidence from the U.S. patent citation data on the dynamics of firms' patent portfolio development, I build a model of innovation incorporating a technology network structure. The model features firms operating in multiple technology sectors and internalizing the spillovers of their own knowledge accumulation to produce patents. Two new insights emerge: The technology network is an important determinant of the patent distribution in different sectors. The growth of patents in each sector is proportional to the Eigenvector Centrality of the technology network. The model is estimated using Simulated Method of Moments and it is capable of reproducing the patent distribution observed in the data.

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, Kevin Lee, Andy Dickerson, James Morley, May Li, Kei Kawakami, Lawrence Uren, Yury Yatsynovich, Lin Shao and the seminar participants at Melbourne University, 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

It has long been recognized that technology network, which refers to the interconnection of knowledge over different technology sectors, is important in affecting research and development (R&D) through the channel of knowledge spillovers.[1] Despite its importance, there is little research on how such technological connections can affect growth at both firm and sector level. This is of particular interest today, given the recurrent debate and practice of industry policies in many countries. For example, The Chinese government proposed to invest more than 161 billion dollars over 10 years to develop the semiconductor industry in 2015. In the Union budget of 2015-2016, the government of India announced plans to provide credit of 127 billion dollars to farmers in order to boost sales in the tractors segment. Remote examples include emerging Asian economies such as Singapore and South Korea, which adopted industry-oriented protection policies during the process of industrialization. The evaluation of policies such as the above will not be precise if we do not have an understanding of the potential spillover effects across industries.

FIGURE 1 showcases an example of a technology network that is of interest to this paper. In the graph, each node is a sector in the economy and each edge implies the existence of technology connection between sectors. The size of each node represents proportion of knowledge flows from a sector to all the other sectors, measured by the backward citation ratio. The thickness of the edge captures the sector-pair knowledge flow. The network as a whole highlights the sectoral connections of an economy. As shown from the graph, there is substantial heterogeneity across sectors in terms of spillovers effects, giving rise to the asymmetric structure of technology network. The main object of the paper is to study the interaction between the technology network and firms' knowledge accumulation and their implications.

This paper proposes a multi-sector model of innovation that takes into account of interdependency of knowledge over different technology sectors. Motivated by the empirical evidence from the patent citation data of U.S. Patent and Trademark Office, I document that knowledge spillovers demonstrate pro-entry and pro-innovation effects. The former facilitates entrants to new sectors while the latter helps speed up the innovation process. I construct a model to incorporate these features by allowing firms to build new patentable ideas upon their past knowledge accumulation from all sectors.[2] This is called internal innovation. Internal innovation allows knowl-

[1] See, among others, David (1990), Rosenberg (1990).

[2] Weitzman (1998) proposes a model where knowledge, through a combinatorics process,

edge spillovers from one sector to not only itself but also other sectors. By contrast, firms may have new ideas that are not directly related to their past ideas, which means that such innovation is independent of their knowledge accumulation. This possibility is captured by random innovation. The key difference between random innovation and internal innovation is that there is a scale effect of knowledge accumulation for the latter while that is not the case for the former.

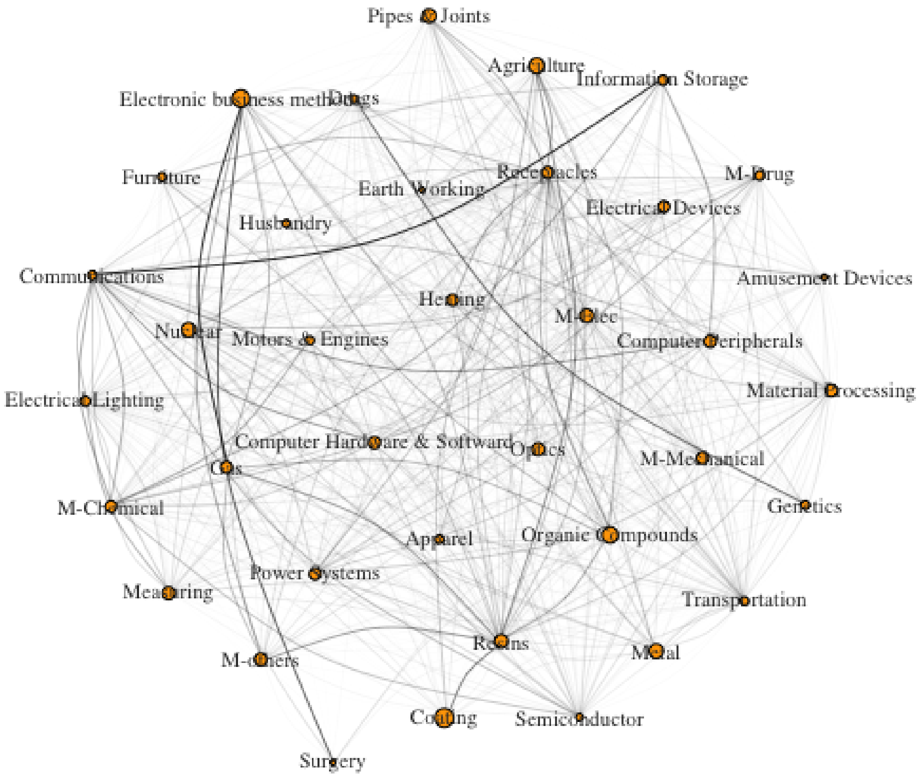


Figure 1
Technology Network

In this framework, firms are characterized by their holding of patents over varied technology sectors, which collectively determine their positions in the technology space. The spillovers of existing ideas contribute to the

can be built upon itself. My notation of internal innovation is similar to his.

discovery of new ideas in the same technology sector as well as other sectors. The technology network is modeled as a matrix with entries capture the relative strength of spillover effects across sectors. An important ingredient of the model is an asymmetric technology network, which enables asymmetric knowledge spillovers across sectors.

Allowing asymmetric knowledge spillovers renders the knowledge accumulation path dependent. Firms accumulate sector-specific knowledge by developing new patents. As they grow, firms differ in terms of the composite of their patent portfolio. Some firms have heavy weight on technology sectors with strong knowledge spillover effects while other firms do not. The difference in firms' patent portfolios, coupled with the asymmetric technology network, leads to the discrepancy in the accumulation of new patents across technology sectors. This feature generates additional challenges to provide any analytical analysis. To proceed, I use mean-field approximation, a methodology commonly used in the network literature,[3] and show that the technology network plays a central role in shaping both the sectoral patent distribution and growth. Specifically, I formally show the shape parameter of sectoral patent distribution is a function of eigenvalue of the technology network. Moreover, the growth of patents in each sector is proportional to the Eigenvector Centrality of the technology network.

I estimate the key parameters of the model to match the total patent distribution, using the Simulated Method of Moment (SMM). The estimates are used to simulate the patent distribution in each individual sector. I show that the simulated patent distribution matches the real distribution observed in the data, a result that lends support to my model.

The paper contributes to the literature in several aspects. First, I document new empirical evidence that demonstrates the importance of the technology network on driving a firm's innovation behavior. Second, I build a model, consistent with the empirical features, that provides a mechanism to understand the connection between technology network and the dynamics of firms' patent portfolio development. Lastly, the paper is the first attempt to explicitly study the relationship between the technology network and the patent distribution as well as the impact of the technology network on knowledge accumulation.

The paper is organized as follows: Section 2 reviews the past literature. Section 3 presents the empirical motivation of this paper. The network innovation model is developed in Section 4 while numerical examples are provided in Section 5 in addition to the test of model performance. Section 6 concludes the paper.

[3]See for example Jackson and Rogers (2007).

2. RELATED LITERATURE

The paper is related to the literature that studies R&D via endogenous innovation.[4] In particular, the model is an extension of Klette and Kortum (2004). In their paper, firms engage in 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. Lentz and Mortensen (2005, 2008) examine a modified version of Klette and Kortum's model by introducing firm heterogeneity and highlight the role played by more productive firms in driving aggregate growth and resource reallocation. Akcigit and Kerr (2010) study the difference between internal and external innovation and its implication for firms' growth. A common feature of the above models is that the realization of new products is undirected. Knowledge accumulation in one sector can result in arrival of products across sectors with equal probability. In contrast, I construct a multi-sector model that focuses on the interdependency of patents over different technology sectors and allow unequal arrival rate of products.

There is a large literature trying to model various networks existing in the economy. Among them, Jackson and Rogers (2007) presents 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 network, the so-called meeting friends of friends. Chaney (2014) provides a network theory of international trade, partly built upon Jackson and Rogers (2007), that features information friction for exporting firms in acquiring new customers overseas. By distinguishing direct search and remote search, he establishes a network structure which represents a preferential attachment model in network theory. The distribution of the number of countries that an exporter has contacts with is a combination of Exponential distribution and Pareto distribution. The similar structure of technology network and social network enables me to use techniques developed in network analysis and derive results linking technology network to firm development dynamics.

This paper is linked to a few attempts to directly incorporate networks into macro analysis. Oberfield (2011) develops a model of a business network through which firms form production chains. Firms produce goods using some techniques they invent and inputs from other firms. Only cost-

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

effective techniques are employed for production. In equilibrium, a star network structure may emerge endogenously when the share of intermediate goods in production is high. In contrast, I take as given the structure of technology network and proceed to study how it shapes firm growth. In a similar application, Cai and Li (2012) study the impact of inter-sectoral knowledge linkages on firms' innovation intensity and the sequence of entry into different industries. In this paper, I examine the impact of technology network on sectoral growth, which is abstracted away from their research.

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. Rossi-Hansberg and Wright (2007) presents a theory of firms' establishment size distribution in an economy where there are many industries, each with numerous identical firms. The accumulation of industry-specific human capital exhibits decreasing return to scale, leading to mean reversion of firm size in equilibrium and therefore a stationary distribution. In a different context, Gabaix (1999) employs a model, underlying which lies a geometric Brownian motion with a reflecting barrier, to explain the city size distribution in United States. In such an environment, the city size distribution demonstrates a heavy Pareto tail resulting from a random growth mechanism.[5] In fact, the tail distribution approaches to Zipf's distribution. This paper shares the common interest in size distribution, but extends the analysis to examine the relationship between technology network and distribution.

This paper is also closely related to Acemoglu et al. (2016). They map the upstream technology network and sectoral patent growth to predict future innovation after 1995 and find strong predictive power at the sector level. The empirical analysis here complements their study by providing additional evidence at the firm level. In addition, I build a model to formalize the idea of knowledge spillovers and test its performance in matching data.

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 the neighbouring firms on the success of 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 the impact of R&D on firms' growth through spillovers using a model that differentiates two types of spillovers: a positive technology spillovers from other research firms and a negative business stealing effect from product

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

market rivals. They employ and extend the measure of technology spillovers from Jaffe (1986) in their empirical parts and conclude that the positive knowledge spillover effects dominate the negative business stealing effects. My paper serves as the complement to the previous studies by considering the internal spillover effect of knowledge. The internal knowledge spillovers are particularly important given that firms can operate in multiple sectors and the knowledge accumulation in each sector depends on each other.

3. MOTIVATION

In this section, I present evidence on the spillover effects of firms' past knowledge accumulation on future innovation. I ask two questions here: Is a firm with a patent portfolio representing a technology mix similar to those embedded in a certain technology sector more likely to enter that sector? Does the past knowledge accumulation stored in patents contribute to the innovation of new patents? These two questions present two different types of spillover effects: the former is what I termed pro-entry effect, which is intended to capture the potential benefit of lowering technological barrier when a firm enters a sector that requires knowledge similar to what the firm has previously accumulated. The latter is pro-innovation effect. The pro-innovation effect tries to study whether past knowledge accumulation helps to speed up innovations in technologically related sectors.

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 to the inventor of each patent, the country the assignees belong to, the date each patent is applied and granted as well as the technology sector each patent belongs to, etc. I use the field classification proposed by Hall et al. (2001), who assign all patents into 37 technology sectors. Details upon how each sector is defined can be found in Hall et al. (2001). Here, for convenience, I label each technology sector by a number from 1 to 37. TABLE A.1 in Appendix shows the correspondence between the numeric label and the original field. I 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 turns out to be critical for us to reveal the technology network of a firm's patent portfolio.

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. I introduce such a measure of technology applicability based on the patent citation. Citation may serve as an indicator of knowledge spillovers. For instance, patent A cites patent B may indicate that knowledge flows from B to A. The measure of technology applicability from sector y to x, $g(y, x)$, is constructed using the following formula:[6]

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

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 targeted technology sector. The measure of average proximity is needed because firms can operate in multiple sectors, and our first measure cannot account for this. I use the variable $Mproximity_{f,t}$ to denote the average proximity, and it is calculated as

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

where $I(y_{f,t} = 1)$ is an indicator variable that takes the value of 1 if the firm f owns at least one patent on technology sector y at time t and zero otherwise. Note that the targeted sector x is excluded in the summation because we want to analyze the cross-sector spillover effects.

As a start point, I run a Probit model as follows:[7]

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

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. I control for whether firm f has previously operated in the same sector as well as the total number of technology sectors firm f operates in at $t-1$. 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, and sector dummies, ζ^x, as well as an idiosyncratic component, $\varepsilon^x_{f,t}$. I follow Wooldridge (2002) to instrument the unobserved firm-sector fixed effect μ^x_f with the time mean of all exogenous variables and

[6] This measure has a nice property that $\sum_x g(y, x) = 1$, which will become useful in the next section.

[7] All regressions in this paper are based on the data between 1990 and 2001. The 2001 end date is chosen to allow for 5-year window for patent reviews. The 1990 begin date allows enough pre-sample data to implement the empirical strategy that will be discussed in details later. Regressions extending the range of data deliver qualitatively same results.

the initial value of dependent variable y_0^x. 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 x contributes to the firm's entrant to x.

TABLE 1
PROBIT REGRESSIONS

	(1)	(2)	(3)	(4)	(5)
$Mproximity_{f,t-1}$	15.33***	14.35***	14.89***	12.82***	15.22***
	(0.16)	(0.12)	(0.13)	(0.15)	(0.24)
$y_{f,t-1}$		2.17***	1.98***	1.91***	1.64***
		(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
$Constant$	-1.67***	-2.09***	-2.27***	-2.64***	-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.

Estimate results are presented in TABLE 1. As shown in column (1) TABLE 1, a firm is more likely to enter a technology sector x when its patent portfolio contains technologies closer to that sector. This is still true after controlling for various factors (column (2) to (4)). The additional control variables behave as we would expect. In particular, a firm that is currently owning patents in x is more likely to continue innovation (column(2)) in the future, and there is a higher probability for a firm with a broad set of patents to expand into another sector (column(3)). The magnitude and significance of the above variables are almost unaffected by adding time and sector dummies to the regression as shown in column (4). The regression results taking into account of unobserved firm-sector effects are provided in column (5) of TABLE 1. After controlling the unobserved firm-sector fixed effects, the magnitude of the coefficient for all controls decline significantly. However, the coefficient for technology distance measure increases compared to column (4), showing the robustness.

One concern with the above approach is that the positive relationship between technology proximity and 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 concern, the Appendix provides Probit analysis for each technology sector. The conclusion drawn is that the pro-entry effect of patent portfolio is indeed observed over all sectors, though there is some heterogeneity in terms of how strong such effect is in each sector. 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 the sector of interest.

Given the positive answer to the first question, I now turn to evaluate how much the past knowledge accumulation contributes to the building of new knowledge. I 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 defined the same way as before. This measure is supposed to capture the spillover effect of the current patent portfolio with respect to sector x. I follow Hall et al. (2005) to construct the patent stock using a perpetual inventory method with a 15% depreciation rate.[8] 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 y at t and $\delta = 0.15$. I then run a Negative Binomial model:

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

where $PatentCounts^x_{f,t}$ represents the number of new patents produced by firm f in the technology sector x from time $t-1$ to time t and $\tilde{\mu}^x_f$ is again the firm-sector fixed effect. $\tilde{\vartheta}_t$ and $\tilde{\zeta}^x$ are sets of time dummies and sector dummies respectively. Due to the nonlinear feature of the Negative Binomial model, I follow Blundell et al. (1999) to use the "pre-sample mean scaling" method to control for fixed effects. The idea is to use pre-sample data of

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

patenting behavior to instrument for unobserved heterogeneity. The long panel of patent data available from USPTO allows me to construct the pre-sample average between 1970 and 1989, which is ideal for the practice. The controls used here are the same as those in (2) except that I include both the dummy to indicate the previous operating status and the lag value of patent counts in the targeted sector.

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)
$YearDummy$	No	No	No	Yes	Yes
$SectorDummy$	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 taken logarithm to make the interpretation convenient. Firm-sector fixed effect is instrumented using the method proposed by Blundell et al. (1999) and shown in column (5). All standard errors are clustered at firm level.

The coefficient of interest is β_1. A positive sign of this coefficient implies the existence of pro-innovation effect. The estimates are shown in TABLE 2. From column (1), it is clear that past patent portfolio does contribute to the production of new patents. Adding the variables that control whether and how many patents firm f owns in the previous period decrease the magnitude of the coefficient (column 2). Nonetheless, inclusion of further controls does not affect it significantly (column (3) to (4)). Finally, in column (5), I demonstrate the results using "pre-sample mean scaling" method to control for the firm-sector fixed effect, which changes neither magnitude nor the significance of the estimate for the weighted patent stock variable. As shown in column (5), past knowledge accumulation has a large effect on the innovation of new patents. 1% increase of weighted patent stock leads to almost 0.4% increase of new patents. [9]

I again leave the sector-by-sector analysis to the Appendix. The results there are consistent with what we found here, and confirm that pro-innovation effect exists over sectors.

4. A NETWORK INNOVATION MODEL

4.1. A Multi-sector Model with Technology Network

The economy consists of a constantly growing population of agents. A fixed proportion of agents become entrepreneurs that come up with patentable ideas and establish new firms. In equilibrium, the number of firms grows at a rate η that matches the growing entrepreneurs, given the initial condition of the economy. Entrepreneurs enter different sectors with different probability. They rely on the existing knowledge stock to come up with new ideas. As a result, the probability to enter a sector is proportional to the knowledge stock in a sector. The innovation activity is modeled as a Poisson process. Two types of innovations occur. Incumbent firms can build new knowledge based on their past knowledge accumulation, which I call *internal* innovation. This type of innovation features the traditional spillover effect of innovation. A distinct feature is that knowledge in different sectors will have varied spillover effects over other sectors depending on their relative technology distance. Alternatively, incumbents can get new ideas by luck, which is independent of their past knowledge accumulation. This type of innovation is termed *random* innovation.

Both types of research activities lead to the discovery of new patents, which possess some value to the owner. I will assume that the value of patents is randomly distributed across firms. There are several ways to interpret the value of a patent. A firm can use its patent to design a new product, and

[9] Also note that in the last column, only observations with pre-sample data are included in the regressions, this is why the number of observations drops significantly.

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, I take the first interpretation. Let's continue to assume that a firm at time t has $n_t(x)$ patents in the technology sector x, and the associated values for these patents are denoted as a vector $\pi_{n_t(x)} = (\{\pi_x(i)\}_{i\in[1,\cdots,n_t(x)]})$. In addition, for a firm with n_t patents, the value of those patents is represented by $\pi_{n_t} = (\{\pi_{n_t(x)}\}_{x\in\mathcal{T}})$, where $n_t = \sum_{x\in\mathcal{T}} n_t(x)$ and $\mathcal{T}$ is the set of all sectors. Note that π_{n_t} is a vector of vectors. For simplicity, I will drop the subscription of time when no ambiguity is caused.

Each firm can be regarded as a collection of research teams, each belongs to a technology sector. Each research team devotes efforts in order to produce new patents. New patents arrive at a rate λ_x.[10] The cost of such research activity is a function of the innovation rate. Moreover, it also depends on the innovation-type-specific productivity, θ_λ.[11] Specifically, the cost function for internal innovation is assumed to take a Cobb-Douglas form:

$$c(\lambda_x) = \lambda_x^\epsilon \theta_\lambda^{1-\epsilon}, \ \epsilon > 1 \tag{4}$$

The assumption $\epsilon > 1$ ensures that the cost function is a convex function of flow innovation rate. Firms with n_x patents and access to the above technology can innovate at an aggregate rate $n_x\lambda_x$. When a patentable idea from internal innovation arrives, it adds a patent to the firm's patent portfolio in certain sector stochastically. The stochastic realization of innovation is governed by a probability function $\tilde{g}(x, y)$, which states the probability of a new patentable idea from the research team at sector x turns out to be a patent at sector y. Note that $\tilde{g}(x, y)$ is the theoretical counterpart of the empirical measure $g(x, y)$ used in the previous section. The matrix spanned by $\tilde{g}(x, y)$ can be interpreted as a technology spillover matrix, the elements of which capture the sector-pair-specific spillover effects. A larger value of $\tilde{g}(x, y)$ implies a larger spillover effect from x to y.

The cost function for random innovation is similar to internal innovation, except an additional fixed cost f. This assumption, as will be clear soon, simplifies our analysis by allowing profit to scale up with the number of patents.[12] The flow cost function of random innovation is:

$$\tilde{c}(\gamma) = \gamma^\sigma \theta_\gamma^{1-\sigma}, \ \sigma > 1 \tag{5}$$

[10] Pay attention that here λ_x is indexed by x, this is because without any assumption, the internal innovation rate could be different across sectors.

[11] This assumption is only to differentiate internal innovation from random innovation.

[12] See Akcigit and Kerr (2010) for using a similar assumption in a different setup.

where θ_γ is the productivity of random innovation. Since random innovation is independent of past innovation, the probability function of the location of new patents is assumed to be constant across sectors.

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

$$
\begin{aligned}
rV(\pi_n) - \dot{V}(\pi_n) = \max_{\gamma, \{\lambda_x\}_{x\in\mathcal{T}}} \Big\{ & \sum_{x\in\mathcal{T}} \sum_{v=1}^{n} \pi_x(v) - \sum_{x\in\mathcal{T}} c(\lambda_x) n_x \\
& + \sum_{x\in\mathcal{T}} \lambda_x n_x (E[V(\pi_{n+1})|x] - V(\pi_n)) \\
& + \gamma(E[V(\pi_{n+1})] - V(\pi_n)) - \tilde{c}(\gamma) - f \Big\} \qquad (6)
\end{aligned}
$$

where $E[V(\pi_{n+1})|x] = \sum_{y\in\mathcal{T}} \tilde{g}(x,y) E[V(\{\pi_{n(s)}\}_{s\neq y} \cup \{\pi_{n(y)+1}\})]$ 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 cost incurred to internal innovation. The second line represents the expected change of the firm's value due to new patents from internal innovation. The total value change is the aggregation of all sectoral change. $\lambda_x n_x$ is the aggregate internal innovation rate in sector x, which is multiplied by the expected value change when one patent is added to the firm's portfolio. It is worthwhile to note that the expected value of the new portfolio $E[V(\pi_{n+1})|x]$ in general depends on which sector the new patent is invented. The reasons are two-fold. Firstly, as discussed previously, research from different sectors could result in new patents across sectors with different probabilities, which means that the expected value of a new patent innovated from different sectors could have varied composite. Secondly, the expected value of patents may not be the same across sectors. To focus on the impact of technology network, I assume that the potential profit from a new patent is the same over sectors, i.e. $E[\pi_x(i)] = \bar{\pi}$, $\forall x, i$. This assumption is justifiable if firms are homogeneous and agents spend the same share of their income on each sector.

The first term in the third line captures the expected value change from random innovation. The difference between random innovation and internal innovation is clear here. While the consequence of internal innovation may depend on x, this is not the case for random innovation. The last two terms in the third line are the variable cost and the fixed cost for random innovation respectively.

The next proposition characterizes the solution to the value function and the associated optimal innovation choices under the assumption that the average profit of a new patent across different technology sectors is the same.

Proposition 1. *Under the assumption* $E[\pi_x(i)] = \bar{\pi}\ \forall x \in \mathcal{T}$, *the value function (14) takes the form:*

$$V(\pi_n) = \sum_{x\in\mathcal{T}} \sum_{v=1}^{n} \frac{\pi_x(v)}{r} + nR \tag{7}$$

where R *is the solution to the equation:*

$$R = \frac{\theta_\lambda}{r}\left(\frac{\frac{\bar{\pi}}{r} + R}{\epsilon}\right)^{\frac{1}{\epsilon-1}}\left(\frac{\bar{\pi}}{r} + R - 1\right) \tag{8}$$

And the optimal innovation choices are given by:

$$\lambda = \theta_\lambda\left(\frac{\frac{\bar{\pi}}{r} + R}{\epsilon}\right)^{\frac{1}{\epsilon-1}}, \quad \gamma = \theta_\gamma\left(\frac{\frac{\bar{\pi}}{r} + R}{\sigma}\right)^{\frac{1}{\sigma-1}} \tag{9}$$

Proof. See Appendix. □

From the proposition, we can see that value function of a firm with a patent portfolio π_n is equal to the discounted additive sum of random profits generated from the whole portfolio plus the option value of research R times the total number of patents. The option value of research is itself a function of the average profit of patents, the productivity of internal innovation, the discounting rate and parameter ϵ. It is easy to verify that the option value of research increases with the average profit of patents as we would expect.

Without heterogeneity across firms, the innovation rate is the same across firms and sectors. The feature allows us to isolate the role of technology spillovers from other factors and facilitates the analysis later.

This section lays out a simple model of firms' innovation process. In the next section, I will explore the reduced form of firms' innovation behavior, and study the impact of technology network on firm size distribution and sectoral growth.

4.2. A Reduced Form Model

As shown in the previous section, ideas for random innovation arrive at a Poisson rate γ, which is constant across firms and over time. When such innovation occurs, it adds a new patent to the firm's patent portfolio, which is located in all possible sectors with equal probability. In contrast, the internal innovation occurs with a mean rate λ. For internal innovation, a realization of patents in a particular technology sector depends on both the past knowledge accumulation as well as the relative technological distance between the realized one and others. The more knowledge a firm acquired in a particular technology sector and its surrounding areas, the more likely it would come up with a new idea in those areas. Formally, the change of a firm f's number of patents in a technology sector x from time t to $t+\Delta t$ for a small time interval Δt is as follows:

$$n^f_{t+\Delta t}(x) - n^f_t(x) = \gamma \Delta t \; Pr(\mathbf{1}[\hat{x}^f_l = x]) + \sum_{y \in \Omega^f_t} n^f_t(y) \lambda \Delta t \; \tilde{g}(y,x) \qquad (10)$$

where $n^f_t(x)$ and $n^f_{t+\Delta t}(x)$ are the number of patents for firm f in technology sector x at time t and $t+\Delta t$ respectively . Ω^f_t is the set of technology sectors that firm f operates in. The condition $\sum_{x \in \mathcal{T}} \tilde{g}(y,x) = 1$ is imposed to satisfy that $\tilde{g}(y,x)$ is a probability function. Equation (10) states that in a small time interval Δt, the change of the number of firm f's patents in a certain technology sector x is a result of patent flows from both random innovation and internal innovation. It is interesting to see that the above equation bears reminiscences of the expression of friendship development in Jackson and Rogers (2007). In their paper, people can acquire new friends either by chance or via their social networks. Here, firms can accumulate new patents due to stochastic reasons or by building new ideas upon knowledge stored in their patent portfolio. The technology network here plays a similar role as the social network in their paper.

A closer look at equation (10) reveals that firms' patent portfolio evolution is path dependent and highly dependent on the structure of technology network. Due to this difficulty, the task of exploring the impact of technology network on firms' development dynamics is left to be done by the numerical analysis in the next section. Here, efforts will be devoted to providing analytical results on the aggregate implications of technology network. First of all, I will focus on the firm size distribution. The dynamics of total number of patents for any firm is independent of $\tilde{g}(y,x)$. To see this, define the total number of patents for firm f at time t as:

$$n^f_t = \sum_{x \in \mathcal{T}} n^f_t(x) \qquad (11)$$

where $n^f_t(x) = 0$ if firm f does not have a patent at x. The evolution of total number of patents is characterized by the following proposition.

Proposition 2. *The dynamics of any firm's patent number is determined by a difference equation that is independent of the firm's position in the technology space. Starting from any time point t, the change of total patents at a very short time interval Δt for firm f is as follows:*

$$n^f_{t+\Delta t} - n^f_t = \gamma \Delta t + \lambda n^f_t \Delta t \qquad (12)$$

Proof. See Appendix. □

Intuitively, the dynamics of total patents in our model is completely determined by the intensities of two types of innovations. **Proposition 2** reveals the relative importance of the two types of innovations on a firm's different development phrases. When a firm is small, it relies more on random innovation simply because there is no existing knowledge stock to build on. As the firm grows and accumulates enough knowledge, the contribution of the random innovation becomes trivial, and the internal innovation dominates the growth of patents.

Proposition 2 provides a means to explore the evolution of the total number of patents, in the remaining part of the section, I will focus on characterizing firm size distribution. For simplicity, I will drop the firm index f anywhere possible if that does not cause any ambiguity.

It is useful to note that equation (12) can be rewritten as a differential equation by dividing both sides with Δt and taking the limit as the time interval goes to 0. Thus we have:

$$n_t^{'} = \gamma + \lambda n_t \tag{13}$$

where $n_t^{'} = \frac{dn_t}{dt}$. This is a standard first-order linear inhomogeneous differential equation that admits a close-form solution given an initial condition $n_0 = 0$.[13] More importantly, the functional form also allows an expression of firm size distribution measured by the total number of patents. This feature is summarized in the following proposition.

Proposition 3. *Given that internal innovation intensity λ is constant across firms, the accumulative distribution of firm size $F(n)$, measured as the number of patents, in the economy is given by:*

$$F(n) = 1 - (1 + \frac{\lambda}{\gamma} n)^{-\frac{\log(1+\eta)}{\lambda}} \tag{14}$$

Proof. See Appendix. □

The intuition of this result can be better comprehended by thinking of two cases that correspond to the two stages of a firm's development. Firstly, consider the case when n is small. This is corresponding to the stage when firms have little knowledge accumulation and thus the random innovation is the key force that drives the innovation procedure. In this case, the firm size distribution $F(n)$ is approximately an exponential distribution $1 - e^{-\frac{\log(1+\eta)}{\gamma} n}$

[13] The specific solution is provided in Appendix.

with rate parameter $\frac{\log(1+\eta)}{\gamma}$.[14] This feature is shown in the *lower* tail of firm size distribution. Notice that the exponential rate parameter $\frac{\log(1+\eta)}{\gamma}$ is independent of the internal innovation rate λ. This is intuitive and due to the fact that the random innovation dominates the internal innovation when n is small, therefore the latter is ignored in approximation.

On the other hand, when n is large, the firm size distribution approaches to a Pareto distribution $1-(\frac{\lambda}{\gamma}n)^{-\frac{\log(1+\eta)}{\lambda}}$ with parameter $\frac{\log(1+\eta)}{\lambda}$.[15] This is the scenario when firms have grown for a while and have accumulated significant amount of knowledge, so they are capable of conducting a large scale internal innovation based on their existing patent stocks. In this case, the relative importance of internal innovation outweighs that of random innovation, thus determining the upper tail distribution of firm sizes.

In this framework, the patent stocks of all firms in the same age cohort grow exponentially. This is immediately clear by examining equation (13) and recognizing that the change rate of total patents is proportional to the current patent stock, ignoring the random innovation. The fact that we end up with a Pareto distribution at the *upper* tail is a consequence of constant growth of firm population. The feature also implies that there is positive relationship between firm age and *average* firm size. In fact, it is easy to see that a higher population growth rate results in a thinner upper tail while a higher internal innovation rate leads to a fatter upper tail.

4.3. Dynamic of Firms' Sectoral Patents

In this subsection, I aim to explore sectoral patent distribution. Based on the previous analysis, it is natural to conjecture that the sectoral patent distribution also follows similar mechanism that generates the total patent distribution. This is evident by examining equation (10), which essentially states that the number of new patents discovered at a specific location of the technology space for a firm is determined by a difference equation of this firm's current patent portfolio across the whole technology space. In order to proceed in our analysis, I use the mean-field approximation that treats the average behavior described in equation (10) as exact.[16] It turns out, under the

[14]This expression can be derived by noting that

$$F(n) = 1 - e^{-\frac{log(1+\eta)}{\lambda}\log(1+\frac{\lambda}{\gamma}n)} \approx 1 - e^{-\frac{n}{\gamma}\log(1+\eta)}$$

where n is small.

[15]This is true since when n is large, $1+\frac{\lambda}{\gamma}n \approx \frac{\lambda}{\gamma}n$.

[16]Note that the innovation process described in equation (10) is in fact random, and more importantly, path dependent. Therefore, it is extremely complicated to analyze the

mean-field approximation, that the upper tail of firm size distribution in each technology sector is a Pareto distribution as well. The following proposition provides the detailed results regarding the characteristics of disaggregated firm size distribution.

Proposition 4. *Under mean-field approximation, the counter accumulative distribution function of an average firm's patent holdings in any technology sector, $s \in \mathcal{T}$, is a Pareto distribution in the upper tail as follows:*

$$1 - F(n^s) \sim (\frac{n^s}{c^s})^{-\frac{\log(1+\eta)}{\tau^s}} \tag{15}$$

where n^s is the number of patents in a technology sector s, c^s is a constant term, which is a function, specific to s, of λ, γ as well as the probability functions $\tilde{g}(y, x)$, and τ^s is an eigenvalue of the matrix G consisting of $\tilde{g}(y, x)$ multiplied by λ, $\forall x, y \in \mathcal{T}$

Proof. See Appendix. □

Let's provide some intuition about the proposition. As discussed before, the evolution of a *typical* firm's patent holding in any technology sector under mean-field approximation is determined by the research activities from both random innovation and internal innovation in all currently operating research lines of the firm. In a continuous version, this relationship can be expressed as $n_t^{'}(x) = \gamma c + \lambda \sum_{y \in \mathcal{T}} n_t(y)\tilde{g}(y, x)$.[17] Therefore, the dynamic of the whole patent portfolio of a firm is essentially an inhomogeneous system of first-order linear differential equations. Over time, the importance of internal innovation outweighs that of random innovation in each technology sector, thus the behavior of this system is eventually governed by the internal innovation.[18] This result may be better explored by examining the two innovation processes separately. The random innovation, acting alone without internal innovation, leads to a constant inflow of new patents at each instant. By contrast, the internal innovation depends on weighted average patent stocks across different technology sectors, with the weights equal to $\tilde{g}(y, x)$. With a constant growing population of firms, the former generates a much thinner tail than the latter, which is basically why the CDF in **Proposition 4** takes the form of a power law distribution in the upper tail.

evolution of this system in a precise way. The mean-field approximation lends us a method to study the average behavior of a cohort by assuming that a firm's innovation in any area of research gets exactly the probability share of all its new patents.

[17]The equation is derived by taking the limit of equation (7).

[18]This is consistent with the behavior of total patent evolution. In fact, since the fat tail of Pareto distribution is preserved under aggregation, there should be no surprise that total patent distribution inherits the characteristics of sector patent distribution.

The sectoral patent distribution is closely related with sectoral growth of patents. Although the internal innovation rate is the same across different technology sectors, this does not mean that the growth of each technology sector is the same. In fact, the sectoral growth rate (measured in terms of number of patents) is proportional to the product of λ and a sector specific factor determined by the sector's position in the technology network. **Proposition 5** summarizes the details.

Proposition 5. *Under mean-field approximation, the growth rate of technology sector s is approximately equal to λq_s, where $q = (\{q_s\}_{s \in \mathcal{T}})$ is the solution to the following system:*

$$q\tilde{G} = q \tag{16}$$

and $\tilde{G} = (\tilde{g}(x, y))$ is the matrix spanned by $\tilde{g}(x, y)$.

Proof. See Appendix. □

The intuition for **Proposition 5** is that as time passes, internal innovation outweighs random innovation in determining the patent growth. Therefore the latter is ignored when calculating the growth rate. In equilibrium, the probability distribution of new patents over different technology sectors is constant and equal to the probability distribution of existing patents. As a result, each technology sector simply grows proportionally to the total patents. It is important to note that the ratio of the number of patents in sector s over total number of patents is determined by the technology network $\tilde{G}$ as a whole. Therefore, one technology sector that receives larger spillovers[19] than the other is not enough to guarantee that the former grows faster than the latter. Only when a sector places heavy weights on spillovers from other fast-growing sectors does it grow quickly itself, provided other things equal. Sectors that grow faster have thicker tails in their patent distribution because, compared to other sectors, they have a bigger fraction of firms that produce larger number of patents.

An alternative interpretation of **Proposition 5** arises from graph theory. Notice that the vector of the sectoral growing factors q shown in equation (16) is actually the leading eigenvector of the technology network $\tilde{G}$, with the corresponding eigenvalue equal to 1.[20] In graph theory, q is the generalized eigenvector centrality of a graph spanned by the nodes representing technology sectors.[21] The eigenvector centrality is a measure of how central a node resides in a network. It assigns high scores to nodes connected with other

[19]The total spillovers received by a sector s is the sum of the s_{th} column of $\tilde{G}$.
[20]1 is the largest eigenvalue to $\tilde{G}$ because $\tilde{G}$ is a probability transition matrix.
[21]See Newman (2004) for details.

high score nodes. In the context of this paper, a technology sector that is connected with other fast growing sectors tends to have a high score. Therefore, **Proposition 5** provides a new mechanism to understand sectoral growth by examining how central a sector's position is in the technology network.

To sum up, in this section, I show that the technology network plays an important place in shaping individual firms' patent portfolio dynamics. The analysis leads to the conclusion that technology network is an important factor of determining the distribution and growth of technology sectors. In the next section, I will use several numerical examples to demonstrate specifically how the picture of a firm's patent portfolio would be different if the knowledge spillover effect does not exist. In addition, I formally test the performance of my model and show that our model is able to produce patterns observed in the data.

5. SIMULATION

The purposes of this section is three-fold. Firstly, I would like to test the validity of mean-field approximation by simulating the innovation process described in the previous section and compare the results to the prediction from the theory (**Proposition 4**). The analytical results of our model heavily depend on the use of mean-field approximation, therefore, it is of vital importance to ensure the validity of this method. Secondly, I intend to examine the interaction of technology network and firm growth through several numerical simulations. Lastly, I would like to estimate the model and compare its performance with the data.

5.1. Validity of Mean-Field Approximation

I use equation (10) to guide the simulation. First of all, I generate two independent Poisson process, one for random innovation and the other for internal innovation. I chose the random and internal innovation rates to be such that $\lambda = 0.05$ and $\gamma = 0.2$. An initial population of 20 firms are simulated for 230 periods. At each period, the number of firms grows at a constant rate of 4%.[22] Next, the probability function $\tilde{g}(y, x)$ dictates the stochastic accumulation of patents for each firm. Since $\tilde{g}(y, x)$ is a theoretical probability function, I use $g(y, x)$ for the purpose of simulation. One difficulty associated with the simulation is that the dynamic of a firm's patent portfolio is history dependent. One firm's patent accumulation in any sector depends

[22] Choices of different parameter values will not change the results qualitatively.

not only on his previous patent stock in that sector but also others because of the spillover effect. As a result, to simulate this process requires storing all information of a firm at each period and recording the stochastic realization according to the probability distribution for each technology sector.

As a comparison, I follow **Proposition 4** to generate the theoretical distribution for each sector. This is done by applying mean-field approximation and let firms get their expected proportions of patents in each sector following the probability matrix spanned by $\tilde{g}(y, x)$. In addition, the number of new patents that a firm innovates per patent at each time period is set to equal the mean, which is equivalent to λ for internal innovation and γ for random innovation.[23]

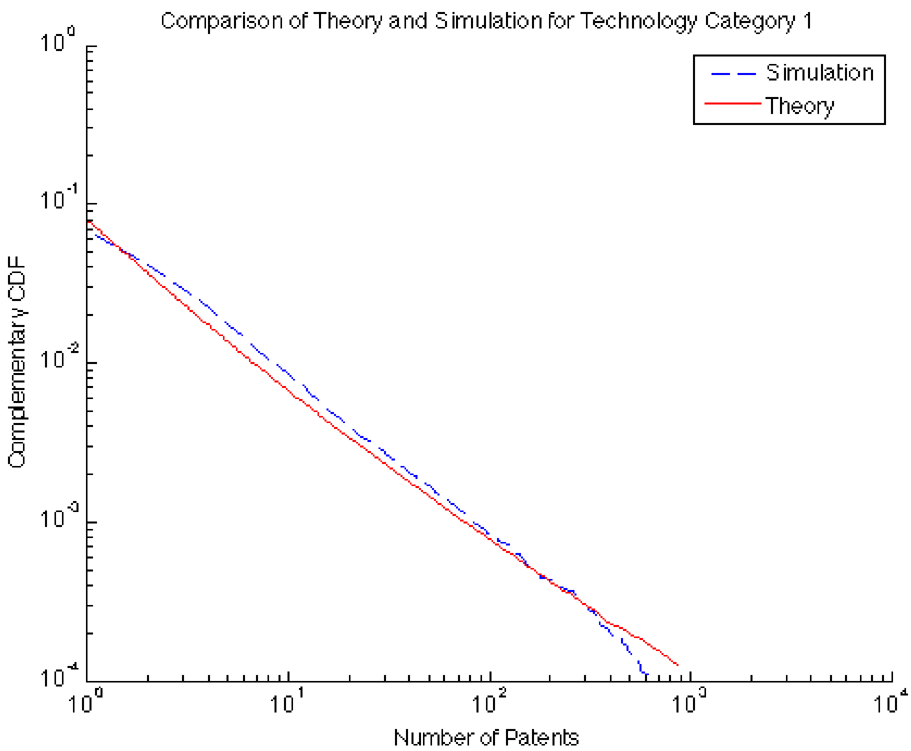


FIGURE 2
THEORY V.S. SIMULATION FOR TECHNOLOGY SECTOR 1

FIGURE 2 demonstrates the results for a randomly chosen sector. The red solid line comes from the theoretical prediction of **Proposition 4** while the blue dash line represent our simulation. The y axis is the complementary

[23]For example, if $\tilde{g}(y, x) = 0.1$ and $\lambda = 0.05$, then a firm with $n_t(x) = 10$ will get 0.05 patent at y next period from her research team at x.

CDF, and the x axis denotes the number of patents. Both axes are drawn at log scale. As shown by the figure, the mean-field approximation performs quite well over a large range of values. Comparisons for other sectors are shown in FIGURE A.1 in Appendix. Again, the mean-field approximation matches the empirical simulation consistently.

FIGURE 3 provides a check for the total distribution by aggregating the data from the previous simulation. As predicted by **Proposition 2**, the distribution of total patents is a mixture of exponential distribution and Pareto distribution. This is clearly reflected by the shape of blue line, which is curly when number of patents is small, but turns out to be a straight line for large number of patents. The prediction using mean-field approximation (red solid line) is smoother than the simulation in the lower tail, but fits the upper tail quite well. The deviation in the lower tail is not surprising given that there is a lot of noises associated with random innovation, which dominates the shape of lower tail. To sum up, our use of mean-field approximation is valid for the purpose of this paper.

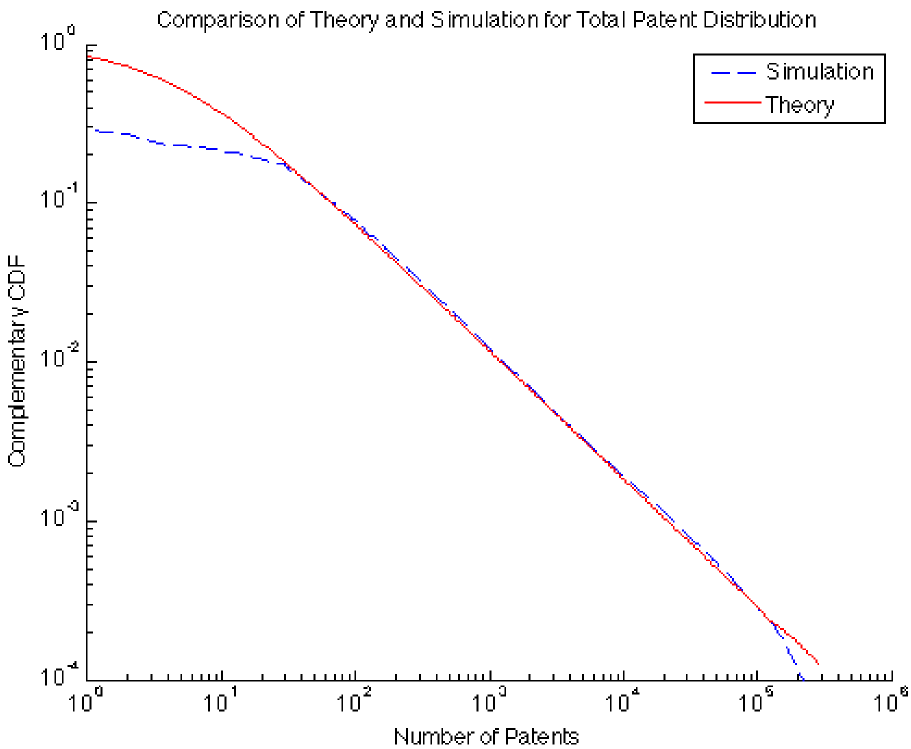
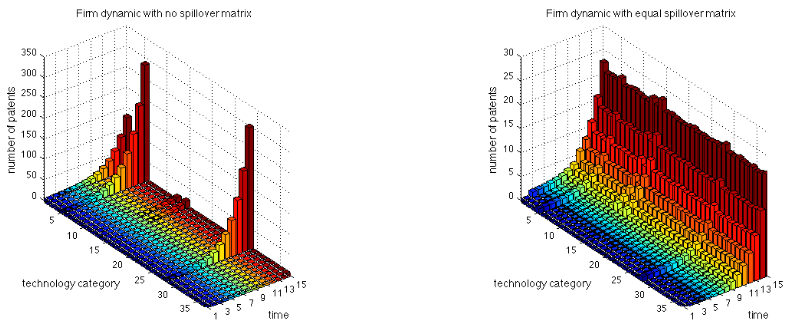


FIGURE 3
THEORY V.S. SIMULATION FOR TOTAL PATENT DISTRIBUTION

5.2. Numerical Analysis

I next proceed to examine how the spillover effect of technology network shapes firms' patent portfolio dynamics. Four exercises are done here: (1) I analyze two counterfactual cases, one with no spillover effect while the other with equal spillover effect, to showcase the impact of different network structures. (2) Given the technology network observed in the data, I simulate firms' dynamics using different initial conditions. (3) I do the same simulation as (2) with a counterfactual technology network to highlight the importance of asymmetric network structure. (4) I compare sectoral patent distributions using the observed technology network and one with equal spillovers.

For exercise (1), the first case corresponds to $\tilde{g}(x, y) = 1$ if $x = y$ and $\tilde{g}(x, y) = 0 \ \forall x \neq y$, while the second case means $\tilde{g}(x, y) = \frac{1}{|\mathcal{T}|} \ \forall x, y$, where $|\mathcal{T}|$ denotes the number of sectors. In order to focus on the network effect, I limit the influence of random innovation on our results by using the same realization of random innovation across firms, i.e. different firms have the same history of random innovation. I simulate a population of 500 firms over 20 periods. To speed up the convergence, both random innovation rate γ and internal innovation λ rate are assumed to be 0.5. FIGURE 4 compares the time dynamic of an *average* firm's patent portfolio under the above two network structures.[24]



(a) (b)

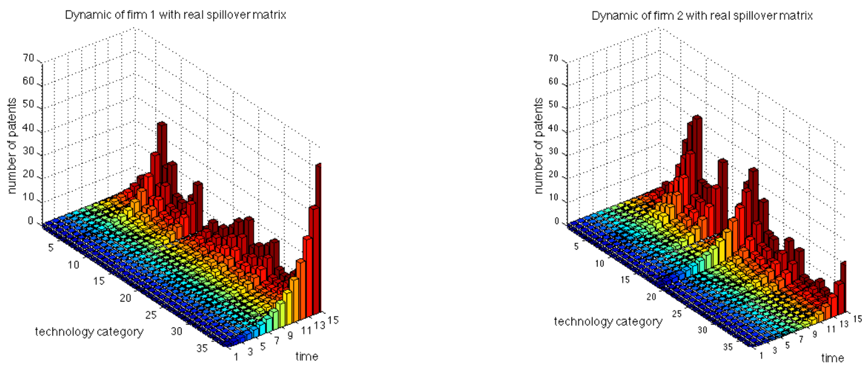
FIGURE 4
COMPARISON OF TWO FIRMS' PATENT PORTFOLIO UNDER DIFFERENT NETWORK STRUCTURES

The left panel depicts the first case with no spillover while the right panel shows the equal spillover case. The most important insight we gain from FIGURE 4 is that the existence of spillover effect significantly broadens

[24] I average the patent stock across 500 firms to represent an average firm.

the range of technology sectors that a firm operates in. Under our extreme assumption of equal spillover, any firm will ultimately operate in every single technology sector as demonstrated by the right panel of the figure.

FIGURE 4 is a stark demonstration of different technology networks. In the next exercise, we will see how firms initially entering different sectors will behave differently due to the asymmetric network structure. I randomly pick up two technology sectors, one for each firm[25] as the first sector they enter. Then I simulate their development over time using the empirical measure of technology network constructed in the motivation section .[26] Again I average out the randomness to better preserve the systematic feature. The results are displayed in FIGURE 5.



(a) (b)

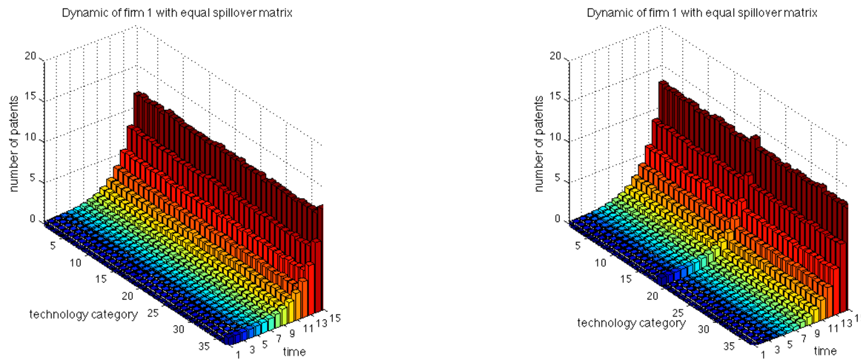
FIGURE 5
COMPARISON OF TWO FIRMS' PATENT PORTFOLIO WITH DIFFERENT INITIAL CONDITIONS

FIGURE 5 confirms that the asymmetric technology network structure strongly influences the development of individual firms. Firms that enter different sectors enjoy different spillover effects of their patent stocks, which determine the probability of entering new sectors. However, this will not be the case if the technology network is symmetric as what I assume for the right panel of FIGURE 4. In exercise (3), I implement a simulation exercise for which everything is identical to (2) except that the spillover matrix spanned by $\tilde{g}(y, x)$ is substituted by the one used for the right panel of FIGURE 4. This is a useful reference case, because it shows that without asymmetric

[25] To avoid the random innovation masks the network effect, I shut down random innovation completely here.

[26] In essence, $\tilde{g}(x, y) = g(x, y)$.

network structure, firm dynamics will not be seen. The comparison is shown in FIGURE 6.



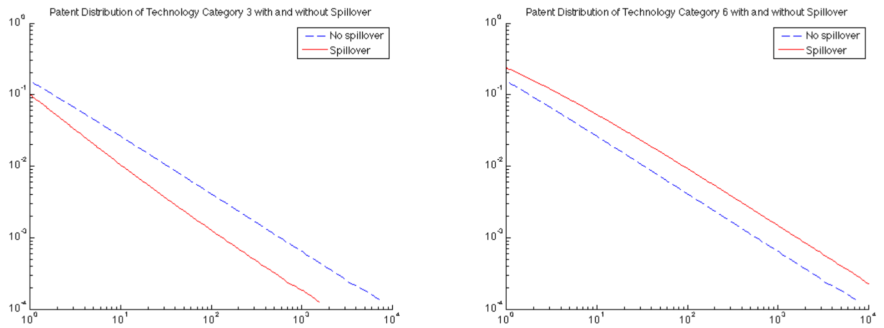
(a) (b)

FIGURE 6

COMPARISON OF TWO FIRMS' PATENT PORTFOLIO USING EQUAL SPILLOVER EFFECT

Not surprisingly, FIGURE 6 shows that firms simply develop a balanced patent portfolio under a symmetric technology network. This is consistent with our expectation.

The last exercise in this subsection tries to explore in depth how the existence of technology network affects the distribution of patents across sectors. This is done by comparing the prediction from our theory (**Proposition 4**) and a counterfactual case where the technology network does not exist. I employ the same code that I use to generate the red line in FIGURE 3 but change $\tilde{g}(x, y)$ to be equal to 1 if $x = y$ and 0 otherwise, which gives us the artificial case of no technology network. Let me provide some intuition for what we expect to see. In the reference case with no technology network, there is no other factor that can affect the thickness of the upper tail of patent distribution, therefore all sectoral distributions should be identical. However, according to **Proposition 4**, the patent distribution of any particular technology sector is determined by two forces: the internal innovation rate λ and the technology network. Since the internal innovation rate is constant across sectors, it is the spillover matrix that is in play here. The existence of the asymmetric technology network enables the patent distribution in each sector to be different. FIGURE 7 demonstrates two examples that are consistent with our expectation. The left panel shows the comparison for sector 3, where the theory predicts a thinner upper tail than the case without spillovers. By contrast, the right panel highlights a case where the theory predicts a thicker upper tail. The results for all the other sectors can be found in FIGURE A.2 in Appendix.



(a) SECTOR 3 (b) SECTOR 6

FIGURE 7
COMPARISON OF SECTORAL PATENT DISTRIBUTION WITH AND WITHOUT SPILLOVER EFFECT

5.3. Simulated Method of Moment Estimation

The paper predicts strong connection between the technology network and sectoral patent distribution. It is important to see whether the model can generate distribution consistent with the data. Towards that purpose, I use the total patent distribution as a target to estimate the parameters using simulated method of moments. The total patent distribution is a good start point because it is independent of the technology network, a feature shown in **Proposition 3**. The estimates of parameters, combined with the technology network, can then be used to simulate sectoral patent distributions, which serve as non-targeted moments. If **Proposition 4** is right, the simulated sectoral patent distributions should be able to match that of real data in the upper tails.

I employ the simulated method of moment (SMM) to estimate the three key parameters in my model: the internal innovation rate λ, the random innovation rate γ and the firm population growth rate η. An initial cohort of 50 firms are simulated for 440 periods given a guess of initial parameter values. The subsequent cohort size is determined by η, the growth rate of firm population. After finishing one round of simulation, I match the CDF of simulated patent distribution to that of data at patent number 2 to 49 with interval of 1 and 50 to 10000 with interval of 25.[27] Each iteration generates

[27]The choice of matching points is designed to capture the overall shape of the distri-

new values of parameters and the procedure continues until convergence. The estimates of parameters for year 1996 to 1999 are shown in TABLE 3.

TABLE 3
PARAMETER VALUES FROM SMM

time	λ	γ	η
1996	0.0268	0.0542	0.0208
1997	0.0279	0.0592	0.0220
1998	0.0262	0.0577	0.0207
1999	0.0276	0.0536	0.0208

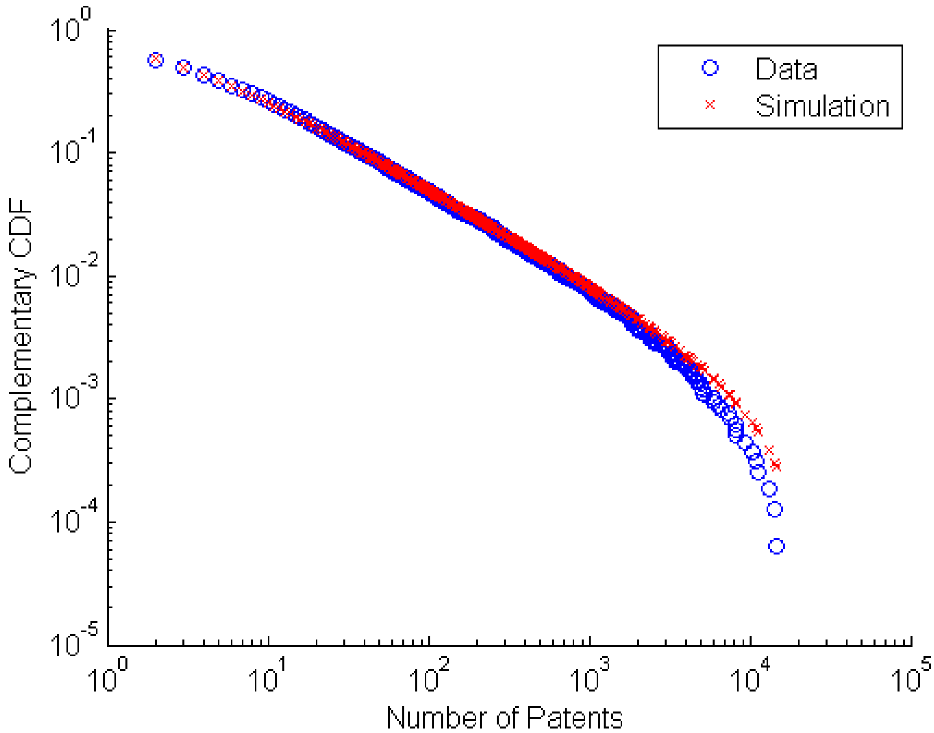


FIGURE 8
TOTAL PATENT DISTRIBUTION: SIMULATION V.S. DATA

The estimate of the internal innovation rate λ is roughly half of the random innovation rate γ. Firm growth rate is just above 2%. The estimates

bution.

vary across years slightly. I use the above estimates to simulate the total patent distribution. FIGURE 8 compares the simulated and real patent distribution for year 1999. It is clear that the model performs well and matches the data closely.

Now we are ready to simulate the sectoral patent distribution. I simulate the process described by equation (10), assuming that each firm gets the expected number of new patents in each sector, which is determined by the two innovation rates λ and γ as well as the technology network $g(x, y)$. To showcase an example of the results, I group together the comparisons of the simulation and the data for two sectors in FIGURE 10. Note that the simulations clearly capture the thickness and the slope of patent distributions in the real data. Results of simulations for all other sectors are shown in FIGURE A.3 in Appendix.

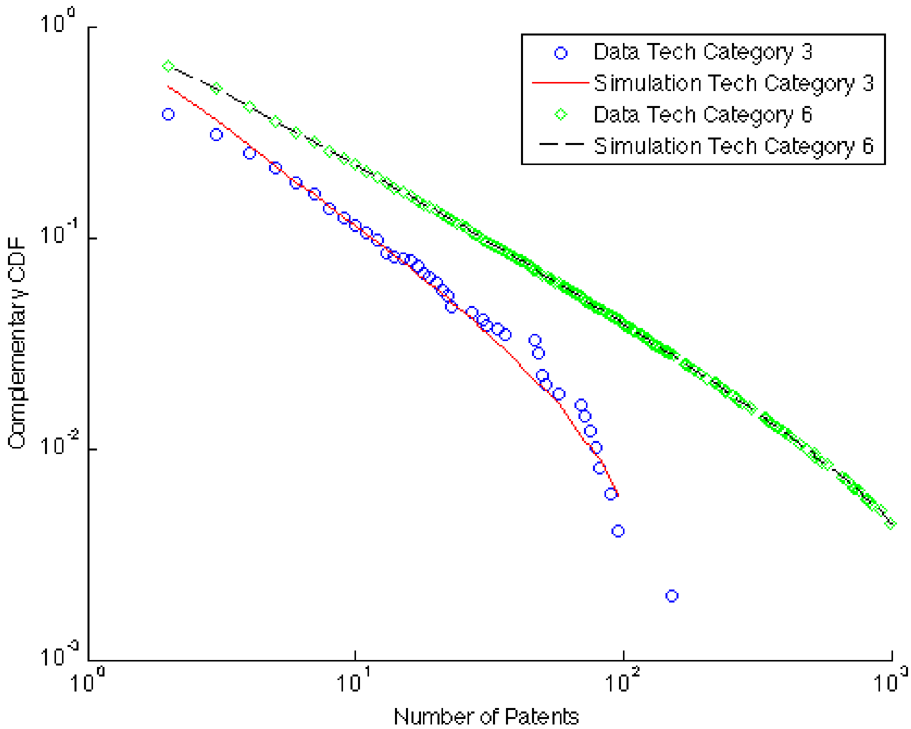


FIGURE 9
PATENT DISTRIBUTION IN TECHNOLOGY SECTOR 3 AND 6: SIMULATION V.S. DATA

In summary, the prediction of the model is consistent with the data at both the aggregate level and sectoral level. The use of mean-field approximation does not affect the validity of the model. Technology networks play a vital role in shaping, not only individual firms' patent portfolio, but also the sectoral distribution of patents. More importantly, the knowledge spillover driven thick tail predicted by the theory fits features of the data reasonably well.

6. CONCLUSION

I construct a dynamic model of innovation on multiple technology sectors, incorporating a network structure of knowledge spillovers. The model is motivated by empirical evidence from the patent citation data collected by U.S. Patent and Trademark Office. The empirical evidence suggests that knowledge spillover is a crucial factor in driving firms to enter new sectors and accumulate new patents. The model features a combination of random innovation and internal innovation, which explains the curvature of total patent distribution at small values as well as the thick upper tails observed in the real data. In addition, the model illustrates how the existence of technology network shapes the dynamics of firms' patent portfolio and provides a new mechanism to understand the relationship between the technology network and sectoral patent distribution as well as patent accumulation. I implement several simulation exercises to show the properties of the model. I use SMM to estimate three key parameters of the model and simulate the patent distribution based on the estimates from SMM. The model's predictions are shown to be a good fit to the data.

The paper is the first attempt to study the relationship between the technology network and the firm size distribution across sectors. The sectoral patent distribution is shown to be a Pareto distribution, with the shape parameters being a function of eigenvalues of a matrix determined by the technology network. This is, to the best of my knowledge, the first time such relationship is provided in the literature. The paper also provides a mechanism to understand the relationship between sectoral patent accumulation and the technology network. The mechanism illustrates how the position of a sector in the technology network can affect the patent accumulation in that sector.

The paper provides a simple framework to analyze the impact of the technology network on the dynamic of firm's knowledge accumulation and growth. Future research can extend the model to incorporate government subsidy or tax to study the welfare effect of such policies. It is easy to imagine that without taking into account of the asymmetric network effect, policy designers may either overestimate or underestimate the policy impact, which depends on the specific positions of firms in the technology space that are subject to the influence of policy. Another direction of future research is to study the impact of the technology network on the diffusion of new technologies. The model in this paper assumes, for simplicity, that spillover from one sector to the other is instantaneous, however, it is not hard to incorporate another parameter to capture the speed of spillovers. Whether the speed of spillovers depends on firms' connectedness in the technology network can be an interesting research question.

Appendix

Proposition 1. *Under the assumption $E[\pi_x(i)] = \bar{\pi}\ \forall x \in \mathcal{T}$, the value function (14) takes the form:*

$$V(\pi_n) = \sum_{x\in\mathcal{T}}\sum_{v=1}^{n}\frac{\pi_x(v)}{r} + nR$$

where R is the solution to the equation:

$$R = \frac{\theta\lambda}{r}(\frac{a\bar{\pi}+R}{\epsilon})^{\frac{1}{\epsilon-1}}(a\bar{\pi}+R-1)$$

And the optimal innovation choices are given by:

$$\lambda = \theta_\lambda(\frac{a\bar{\pi}+R}{\epsilon})^{\frac{1}{\epsilon-1}}, \quad \gamma = \theta_\gamma(\frac{a\bar{\pi}+R}{\sigma})^{\frac{1}{\sigma-1}}$$

Proof. Guess the value function $V(\pi_n) = \sum_{x\in\mathcal{T}}\sum_{v=1}^{n} a\pi_x(v) + nR$. Substitute the guessing value function form into equation (10), then we have:

$$\begin{aligned} ra\sum_{x\in\mathcal{T}}\sum_{v=1}^{n}\pi_x(v) + rnR = \max\{ & \sum_{x\in\mathcal{T}}\sum_{v=1}^{n}\pi_x(v) - \sum_{x\in\mathcal{T}}c(\lambda_x)n_x + \sum_{x\in\mathcal{T}}\lambda_x n_x a\bar{\pi} \\ & + \sum_{x\in\mathcal{T}}\lambda_x n_x R + \gamma a\bar{\pi} + \gamma R - \tilde{c}(\gamma) - f\} \end{aligned}$$

The above equation holds if and only if:

$$\begin{aligned} a &= \frac{1}{r} \\ rnR &= \max_\lambda\{\lambda n(a\bar{\pi}+R) - c(\lambda)n\} \\ f &= \max_\gamma\{\gamma(a\bar{\pi}+R)\} \end{aligned}$$

The second equation could be further simplified to be:

$$rR = \max_\lambda\{\lambda(a\bar{\pi}+R) - c(\lambda)\}$$

Note that there is no sectoral subscript for the internal innovation rate here because given our assumption that an additional patent brings in on average the same profit regardless which technology sector it belongs to, the internal innovation rate is supposed to be the same across sectors. Solve the two optimization problems using $a = \frac{1}{r}$ to get:

$$c'(\lambda) = \frac{\bar{\pi}}{r} + R$$
$$\tilde{c}(\gamma) = \frac{\bar{\pi}}{r} + R$$

It is no surprise that the two first-order conditions tell us that the marginal benefits of both types of innovation are exactly identical to each other and equal to the corresponding marginal costs. Substitute equations (8) and (9) into the above two equations and solve for λ and γ:

$$\lambda = \theta_\lambda(\frac{\frac{\bar{\pi}}{r} + R}{\epsilon})^{\frac{1}{\epsilon-1}}$$
$$\gamma = \theta_\gamma(\frac{\frac{\bar{\pi}}{r} + R}{\sigma})^{\frac{1}{\sigma-1}}$$

which are as desired in the proposition. Substitute the expression for λ back to the optimization problem:

$$\begin{aligned}
rR &= \theta_\lambda \epsilon^{-\frac{1}{\epsilon-1}}(a\bar{\pi} + R)^{\frac{\epsilon}{\epsilon-1}} - (\frac{a\bar{\pi} + R}{\epsilon})^{\frac{1}{\epsilon-1}}\theta_\lambda \\
&= \theta_\lambda \epsilon^{-\frac{1}{\epsilon-1}}[(a\bar{\pi} + R)^{\frac{\epsilon}{\epsilon-1}} - (a\bar{\pi} + R)^{\frac{1}{\epsilon-1}}] \\
&= \theta_\lambda(\frac{a\bar{\pi} + R}{\epsilon})^{\frac{1}{\epsilon-1}}(a\bar{\pi} + R - 1) \\
&= \theta_\lambda(\frac{\frac{\bar{\pi}}{r} + R}{\epsilon})^{\frac{1}{\epsilon-1}}(\frac{\bar{\pi}}{r} + R - 1)
\end{aligned}$$

$\square$

Proposition 2. *The dynamics of any firm's patent number is determined by a difference equation that is independent of the firm's position in the technology space. Starting from any time point t, the change of total patents at a very short time interval Δt for firm f is as follows:*

$$n^f_{t+\Delta t} - n^f_t = \gamma \Delta t + \lambda n^f_t \Delta t$$

Proof.

$$\begin{aligned}
n^f_{t+\Delta t} - n^f_t &= \sum_{x\in\mathcal{T}} (n^f_{t+\Delta t}(x) - n^f_t(x)) \\
&= \gamma\Delta t \sum_{x\in\mathcal{T}} Pr(\mathbf{1}[\hat{x}^f_l = x]) + \sum_{x\in\mathcal{T}} \sum_{y\in\Omega^f_t} n^f_t(y)\lambda\Delta t\ \tilde{g}(y,x) \\
&= \gamma\Delta t + \lambda\Delta t \sum_{y\in\Omega^f_t} \sum_{x\in\mathcal{T}} \tilde{g}(y,x) n^f_t(y) \\
&= \gamma\Delta t + \lambda\Delta t \sum_{y\in\Omega^t_f} n^f_t(y) \\
&= \gamma\Delta t + \lambda\Delta t n^f_t
\end{aligned}$$

The first equality uses the fact that $n^f_t = \sum_{x\in\mathcal{T}} n^f_t(x)$. The third equality exploits the fact that switching the summation does not change the equality. The fourth equality is due to the assumption that $\sum_{x\in\mathcal{T}} \tilde{g}(y,x) = 1$. Note that in the first equality, I use $x \in \mathcal{T}$ instead of $x \in \Omega^f_t$, this is because when new innovation occurs, it could result in new patents outside the firm's current operating technology sets Ω^f_t. □

Proposition 3. *Given that internal innovation intensity λ is constant across firms, the accumulative distribution of firm size $F(n)$, measured as the number of patents, in the economy is given by:*

$$F(n) = 1 - (1 + \frac{\lambda}{\gamma} n)^{-\frac{\log(1+\eta)}{\lambda}}$$

Proof. First of all, it is easy to verify that equation (6) admits a solution as follows:

$$n_t = \frac{\gamma}{\lambda}(e^{\lambda t} - 1)$$

given $n_0 = 0$. Invert the above equation to express time t as a function of n:

$$t = \frac{\log(1 + \frac{\lambda}{\gamma} n)}{\lambda}.$$

In an economy with population of firms growing at η, the proportion of firms with number of patents more than n is pinned down by:

$$1 - F(n) = (1 + \eta)^{-\frac{\log(1+\frac{\lambda}{\gamma} n)}{\lambda}}$$

The above equation can be rearranged to get:

$$F(n) = 1 - (1 + \frac{\lambda}{\gamma} n)^{-\frac{\log(1+\eta)}{\lambda}}$$

as required in the proposition. $\square$

Proposition 4. *Under mean-field approximation, the counter accumulative distribution function of firm's patent holdings in any technology sector, $s \in \mathcal{T}$ is a Pareto distribution in the upper tail as follows:*

$$1 - F(n^s) \sim (\frac{n^s}{c^s})^{-\frac{\log(1+\eta)}{\tau^s}}$$

where n^s is the number of patents in a technology sector s, c^s is a constant term, which is a function, specific to s, of λ, γ as well as the probability functions $\tilde{g}(y, x)$ and τ^s is an eigenvalue of the matrix G consisting of $\tilde{g}(y, x)$ multiplied by λ

Proof. Rewrite equation (7) in continuous form:

$$dn_t(x) = \gamma c + \lambda \sum_{y \in \mathcal{T}} n_t(y)\tilde{g}(y, x), \quad \forall x \in \mathcal{T}$$

Let $N = |\mathcal{T}|$ be the number of elements in set $\mathcal{T}$, $n_t = (n_t(s_1), n_t(s_2), \cdots, n_t(s_N))$, $\Gamma = (\gamma c, \cdots, \gamma c)'$ and a matrix G defined as:

$$G = \begin{bmatrix} \lambda g(s_1, s_1) & \lambda g(s_1, s_2) & \cdots & \lambda g(s_1, s_N) \\ \lambda g(s_2, s_1) & \lambda g(s_2, s_2) & \cdots & \lambda g(s_2, s_N) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda g(s_N, s_1) & \lambda g(s_N, s_2) & \cdots & \lambda g(s_N, s_N) \end{bmatrix}$$

With the above notations, I can rewrite the differential equations system in a compact matrix form:

$$dn_t = G' n_t + \Gamma$$

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

$$n_t = G'^{-1} e^{G' t} \Gamma - G'^{-1} \Gamma$$

Assume that G has distinct real eigenpairs $(\tau_1, \mathbf{v}_1), (\tau_2, \mathbf{v}_2), \cdots, (\tau_N, \mathbf{v}_N)$, [28] then I can decompose $e^{G' t}$ to be as follows:

$$e^{G' t} = P e^{At} P^{-1}$$

[28] This is a technical assumption that greatly simplifies our analysis.

where $P = \mathbf{aug}(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N)$, and

$$A = \begin{bmatrix} \tau_1, & 0, & \cdots & 0 \\ 0, & \tau_2, & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \cdots & \cdots & \cdots & \tau_N \end{bmatrix}$$

Substitute the transformation of $e^{G' t}$ into our solution to the system of differential equations,

$$n_t = G'^{-1} P e^{At} P^{-1} \Gamma - G'^{-1} \Gamma = B e^{At} \mathbf{d} - \mathbf{f}$$

where $B = G'^{-1} P$, $\mathbf{d} = P^{-1}\Gamma$ and $\mathbf{f} = G'^{-1}\Gamma$. It is convenient to work with a specific equation of the system:

$$n_t(s) = b_{s_1} d_1 e^{\tau_1 t} + b_{s_2} d_2 e^{\tau_2 t} + \cdots + b_{s_N} d_N e^{\tau_N t} - f_s$$

where b_{s_i} is the s row i column element of B, and d_i and f_i are the i_{th} element of $\mathbf{d}$ and $\mathbf{f}$ respectively. Let's define τ^s to be such that:

$$\tau^s = \max(\{\tau_i\}_i | b_{s_i} d_i > 0)$$

then I have:

$$\begin{aligned} \lim_{t\to\infty} n_t(s) &= \lim_{t\to\infty} e^{\tau^s t}[b_{s_1} d_1 e^{(\tau_1 - \tau^s)t} + \cdots + b_{\tau^s} d_{\tau^s} + \cdots + b_{s_N} d_N e^{(\tau_N - \tau^s)t} - f e^{-\tau^s t}] \\ &= b_{\tau^s} d_{\tau^s} e^{\tau^s t} \\ &= c^s e^{\tau^s t} \end{aligned}$$

where $c^s = b_{s_{max}} d_{max}, \forall s \in \mathcal{T}$. Such an eigenvalue must exist for each sector, otherwise $n_t(s)$ will either converge to negative infinity or zero, both are impossible here. Nonetheless, they may not be unique. Treating the relationship as exact for large t and invert the above equation:

$$t = \frac{1}{\tau^s} \log \frac{n^s}{c^s}$$

In an economy with population of firms growing at η, the proportion of firms with number of patents greater than n^s approaches to

$$1 - F(n^s) \sim (\frac{n^s}{c^s})^{-\frac{\log(1+\eta)}{\tau^s}}$$

in the upper tail. $\square$

Proposition 5. *Under mean-field approximation, the growth rate of technology sector s in the long run is approximately equal to λq_s, where $q = (\{q_s\}_{s\in\mathcal{T}})$ is the solution to the following system:*

$$q\tilde{G} = q \tag{17}$$

and $\tilde{G} = (\tilde{g}(x, y))$ is the matrix spanned by $\tilde{g}(x, y)$.

Proof. Suppose at the beginning of the date, the distribution of patents across technology sectors is denoted by a vector q_0, elements of which represents the proportion of the number of patents in the corresponding sectors over the total number of patents. Recall that the total number of patents grows at roughly the rate of internal innovation (ignoring random innovation because internal innovation dominates), then one way to think of the sectoral growth is to calculate the probability of allocating every newly arrived patent into the particular technology sector. If such a probability exists and is constant, then we are done. Towards that purpose, we need some notation. Assume q_t denotes the probability distribution of patents across technology sectors at time t, and

$$\tilde{G} = \begin{bmatrix} g(1,1) & g(1,2) & \cdots & g(1,N) \\ g(2,1) & g(2,2) & \cdots & g(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ g(N,1) & g(N,2) & \cdots & g(N,N) \end{bmatrix}$$

then we have:

$$\begin{aligned} q_1 &= \frac{1}{1+\lambda} q_0 + \frac{\lambda}{1+\lambda} q_0 \tilde{G} \\ q_2 &= \frac{1}{1+\lambda} q_1 + \frac{\lambda}{1+\lambda} q_1 \tilde{G} \\ \vdots & \quad \vdots \\ q_t &= \frac{1}{1+\lambda} q_{t-1} + \frac{\lambda}{1+\lambda} q_{t-1} \tilde{G} \end{aligned}$$

The above equations state that the probability distribution of sectoral patents at time t is equal to weighted sum of the probability distribution at time $t-1$ and the probability distribution of newly arrived patents. $q_{t-1}\tilde{G}$ represents the unconditional probability distribution of patents from the innovation at time $t-1$. Since each existing patent give rises to, on average, λ new patents, the weight for the new patents is $\frac{\lambda}{1+\lambda}$. Rearrange the equations and take the time limit:

$$\begin{aligned}
q_t &= q_{t-1}\left(\frac{1}{1+\lambda}I + \frac{\lambda}{1+\lambda}\tilde{G}\right) \\
\lim_{t\to\infty} q_t &= \lim_{t\to\infty} q_{t-1}\left(\frac{1}{1+\lambda}I + \frac{\lambda}{1+\lambda}\tilde{G}\right) \\
q &= q\left(\frac{1}{1+\lambda}I + \frac{\lambda}{1+\lambda}\tilde{G}\right) \\
q &= q\tilde{G}
\end{aligned}$$

where I is identity matrix, and $q = \lim_{t\to\infty} q_t$. q is the stable probability distribution of patents over different technology sectors. Therefore, the growth rate of technology sector s in the long run is equal to λq_s as shown 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)
$Mproximity_{f,t-1}$	46.53***	24.53***	42.15***	19.95***	19.08***	16.22***	15.53***	8.16***
	-9.41	-2.06	-7.62	-1.14	-1.01	-0.5	-0.67	-0.39
$y_{f,t-1}$	1.84***	1.28***	1.56***	1.92***	1.67***	1.57***	1.64***	1.31***
	-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	0.01	0.01	-0.02***	0	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-sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-3.03***	-2.63***	-2.96***	-2.65***	-2.72***	-1.88***	-2.39***	-2.50***
	-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)
$Mproximity_{f,t-1}$	26.34***	20.40***	21.18***	8.74***	13.16***	144.61***	30.94***	26.02***
	-1.76	-1.21	-1.73	-0.42	-0.91	-25.68	-2.89	-1.57
$y_{f,t-1}$	1.56***	1.76***	1.36***	2.26***	2.14***	2.21***	2.05***	1.51***
	-0.06	-0.06	-0.06	-0.04	-0.05	-0.09	-0.08	-0.04
$N_{f,t-1}$	0.01	0	0.01**	-0.01*	-0.01	0.03***	0.01*	0
	-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-sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Constant$	-3.03***	-2.91***	-3.30***	-2.49***	-2.23***	-3.38***	-2.81***	-2.51***
	-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)
$Mproximity_{f,t-1}$	42.93***	15.26***	25.68***	21.49***	16.31***	53.36***	28.77***	47.90***
	-4.35	-1.21	-2.68	-1.35	-1.23	-3.05	-1.17	-2.4
$y_{f,t-1}$	1.59***	1.21***	1.36***	1.46***	1.76***	1.41***	1.46***	1.40***
	-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***	0.01**	0.03***	0.01**	0	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
$Constant$	-2.70***	-2.19***	-2.48***	-2.50***	-2.99***	-2.54***	-2.01***	-2.28***
	-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)
$Mproximity_{f,t-1}$	27.56***	25.16***	49.67***	37.08***	44.27***	184.59***	77.48***	21.62***
	-1.82	-2.57	-3.02	-1.5	-3.1	-17.04	-7.15	-2.43
$y_{f,t-1}$	1.64***	1.54***	1.97***	1.56***	2.08***	2.33***	1.86***	2.15***
	-0.05	-0.06	-0.05	-0.04	-0.06	-0.13	-0.07	-0.07
$N_{f,t-1}$	0	0.03***	-0.01	-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
$Constant$	-2.52***	-2.90***	-2.63***	-1.98***	-2.55***	-2.99***	-2.72***	-2.65***
	-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)
$Mproximity_{f,t-1}$	43.26***	84.72***	95.31***	46.15***	16.21***
	-2.72	-8.25	-8.95	-2.84	-0.51
$y_{f,t-1}$	1.78***	1.62***	1.41***	1.64***	1.43***
	-0.05	-0.07	-0.06	-0.05	-0.03
$N_{f,t-1}$	-0.01	0.01**	0.01**	-0.01**	-0.02***
	-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
$Constant$	-2.48***	-2.74***	-2.87***	-2.32***	-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-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

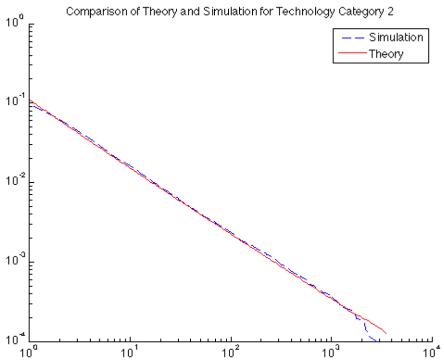
TABLE A.3 Continued

	(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

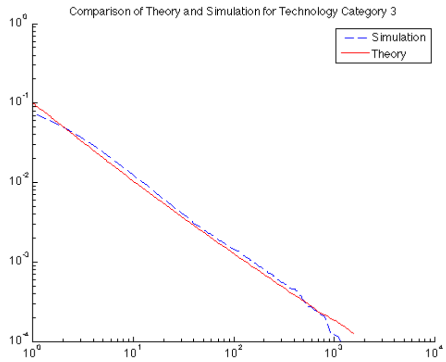
TABLE A.3 Continued

	(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

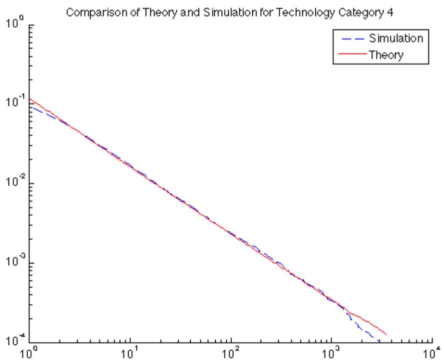
The results in this table are based on the estimation of Negative Binomial regression (3) for each sector. Note that there is no sector dummies here since the regression itself is sector specific.



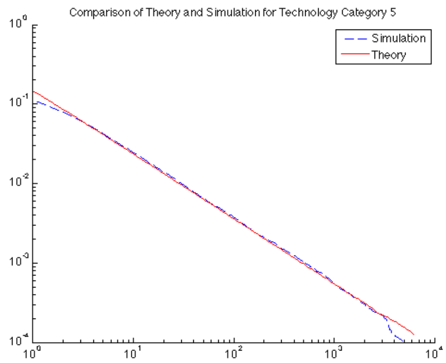
(1) Sector 2



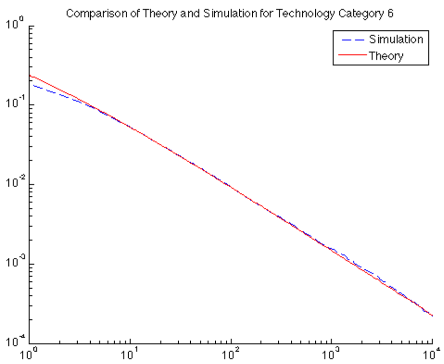
(2) Sector 3



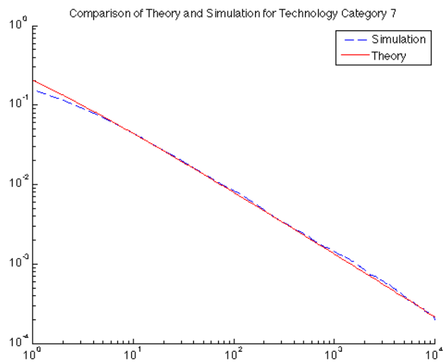
(3) Sector 4



(4) Sector 5

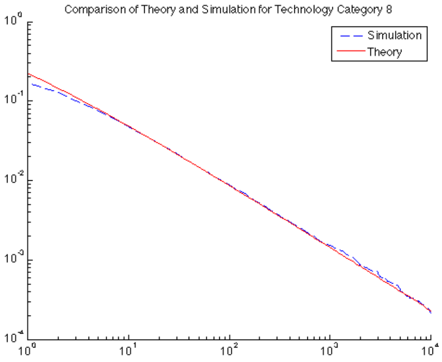


(5) Sector 6

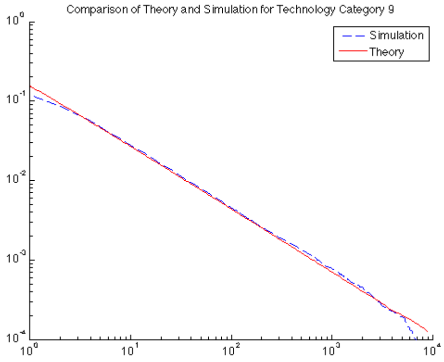


(6) Sector 7

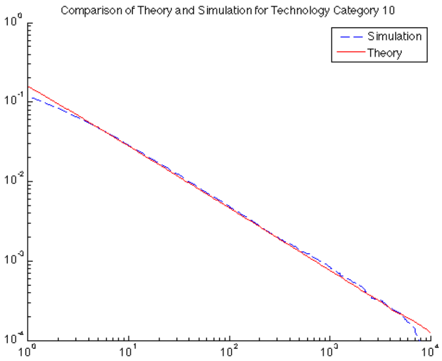
Figure A.1
Theory v.s. simulation for technology sectors across sectors



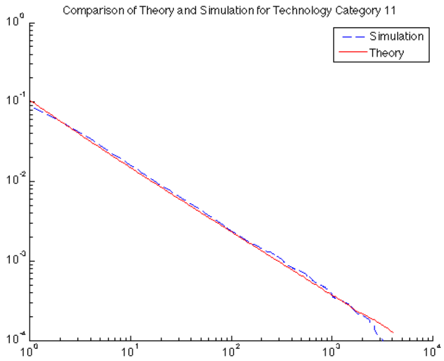
(7) Sector 8



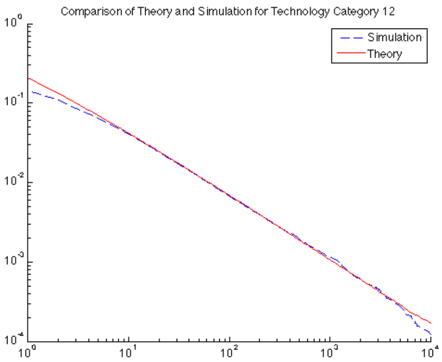
(8) Sector 9



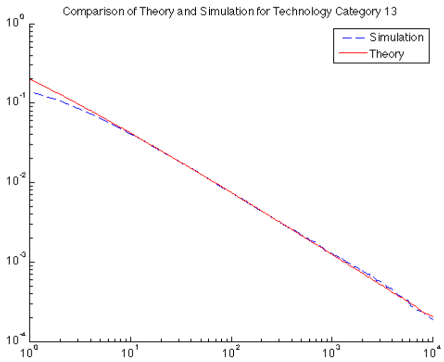
(9) Sector 10



(10) Sector 11

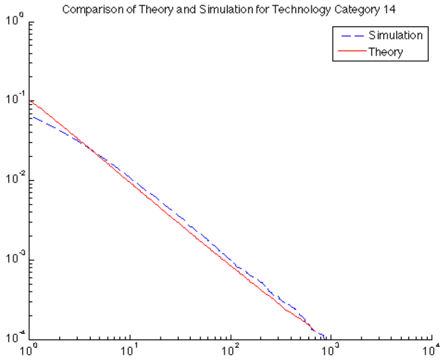


(11) Sector 12

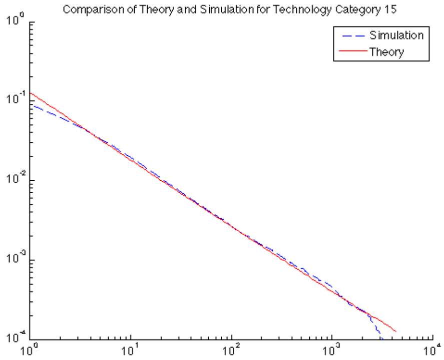


(12) Sector 13

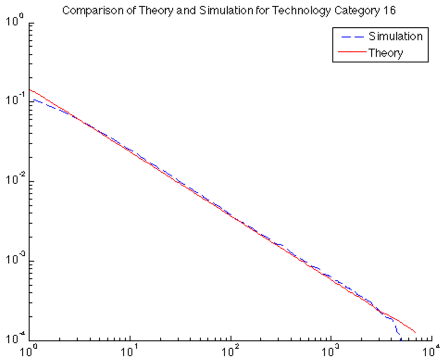
Figure A.1
Continued



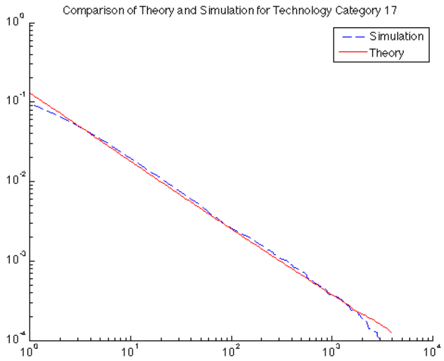
(13) Sector 14



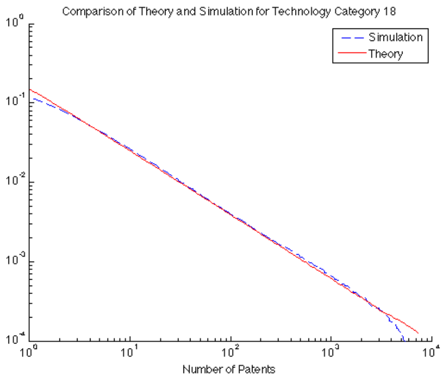
(14) Sector 15



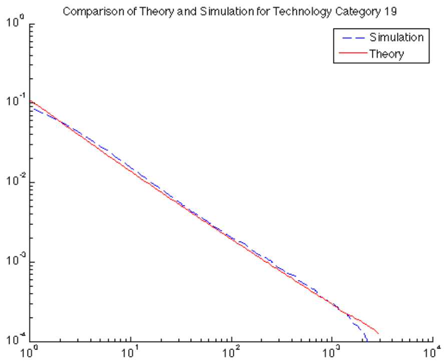
(15) Sector 16



(16) Sector 17

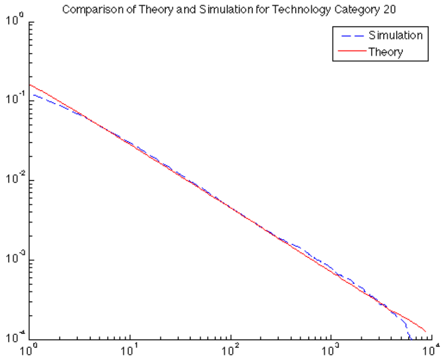


(17) Sector 18

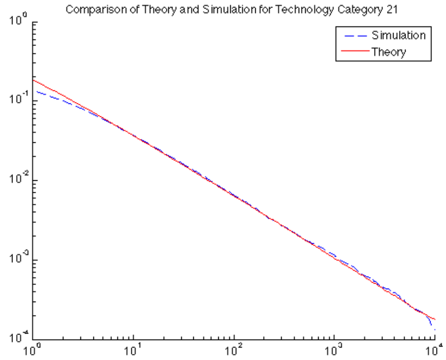


(18) Sector 19

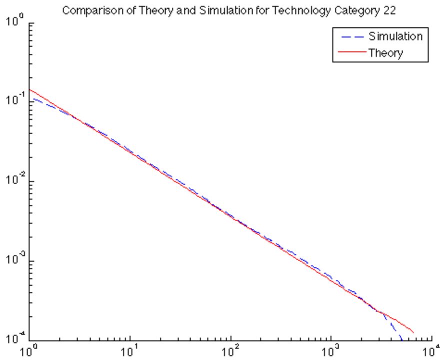
Figure A.1
Continued



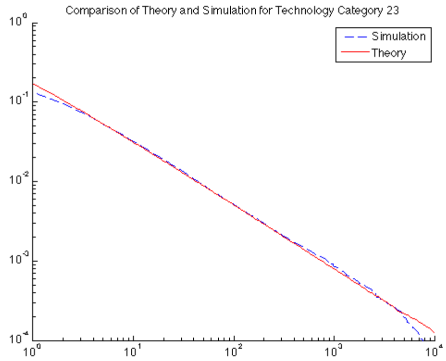
(19) Sector 20



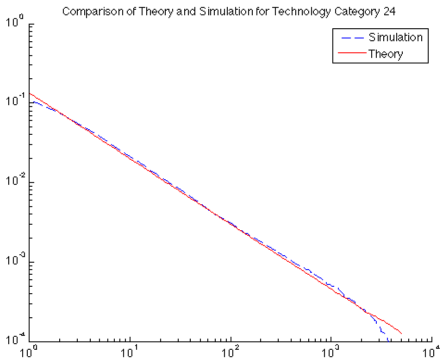
(20) Sector 21



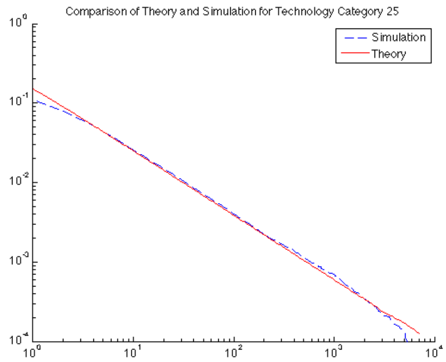
(21) Sector 22



(22) Sector 23

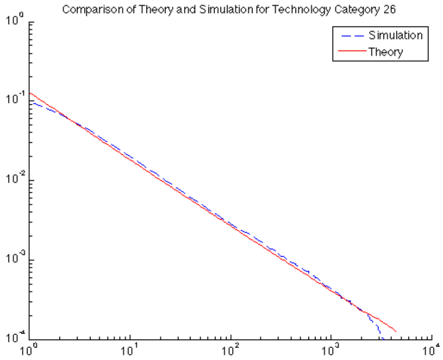


(23) Sector 24

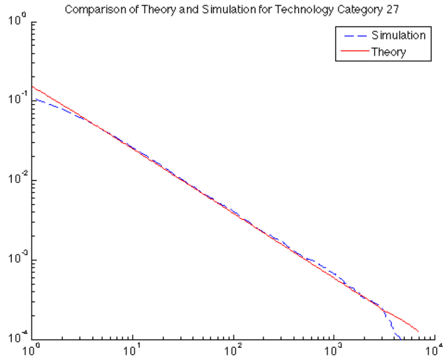


(24) Sector 25

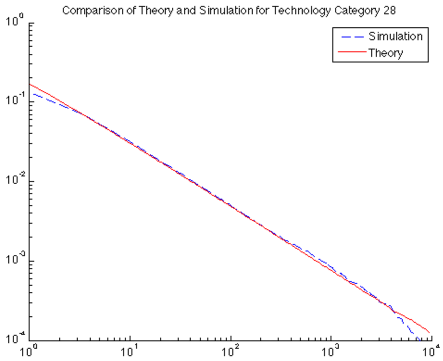
Figure A.1
Continued



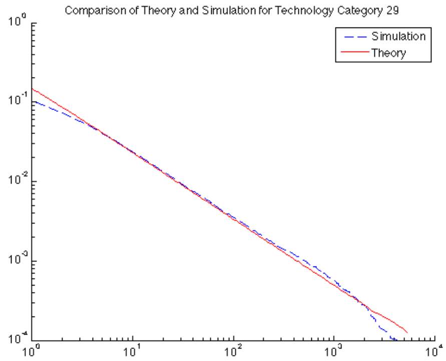
(25) Sector 26



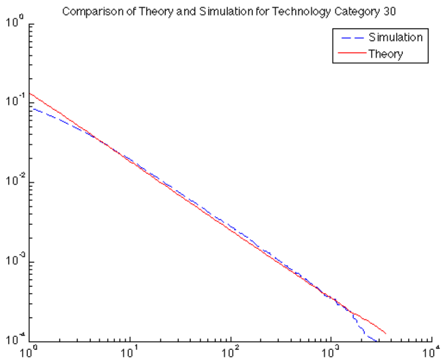
(26) Sector 27



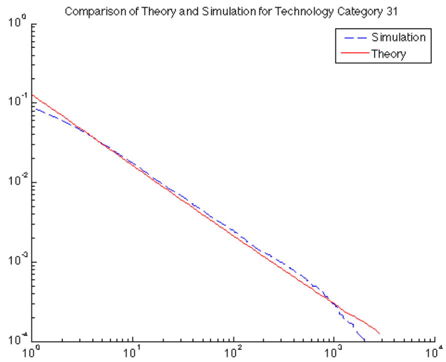
(27) Sector 28



(28) Sector 29

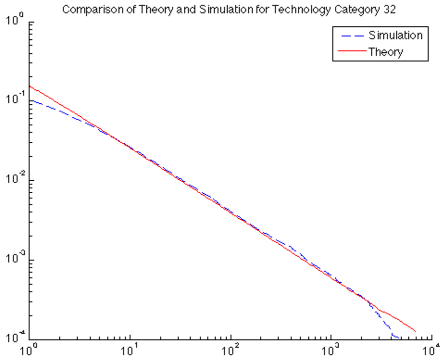


(29) Sector 30

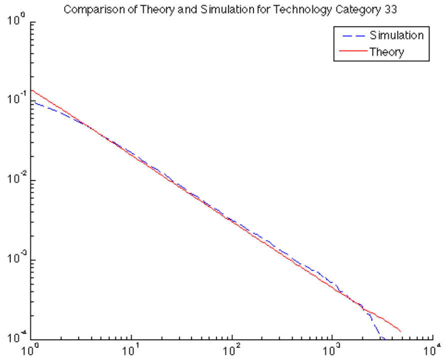


(30) Sector 31

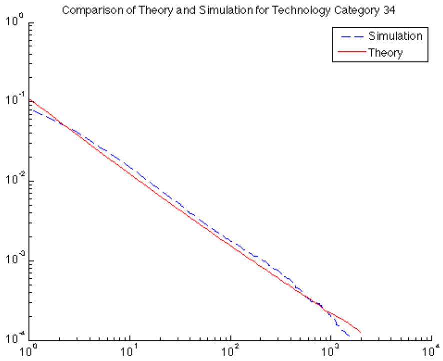
Figure A.1
Continued



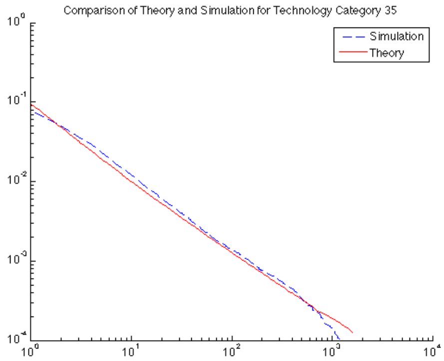
(31) Sector 32



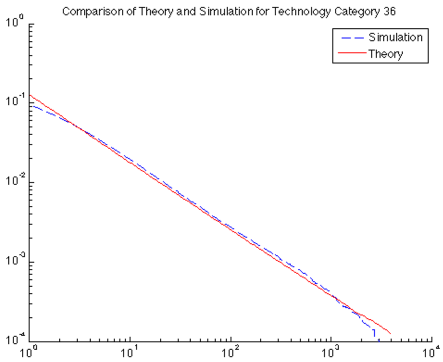
(32) Sector 33



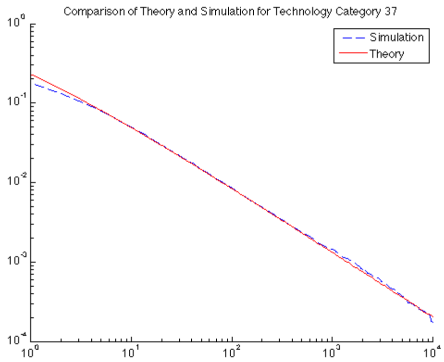
(33) Sector 34



(34) Sector 35

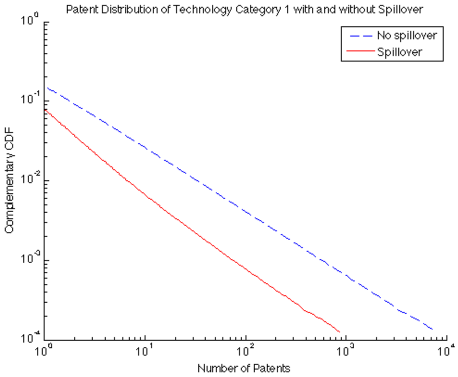


(35) Sector 36

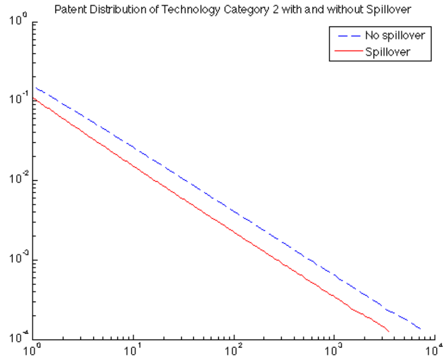


(36) Sector 37

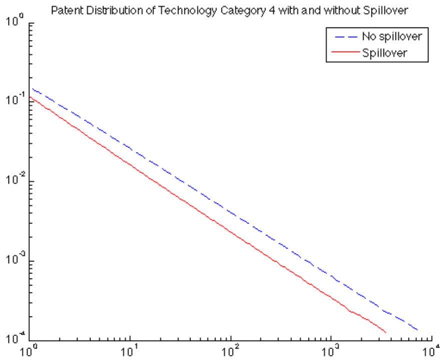
Figure A.1
Continued



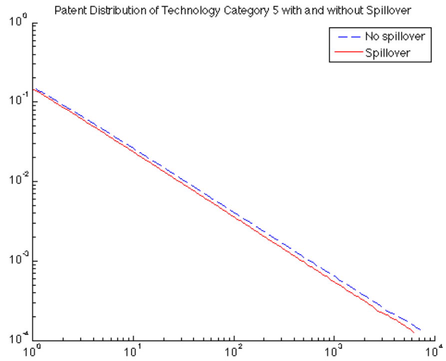
(1) Sector 1



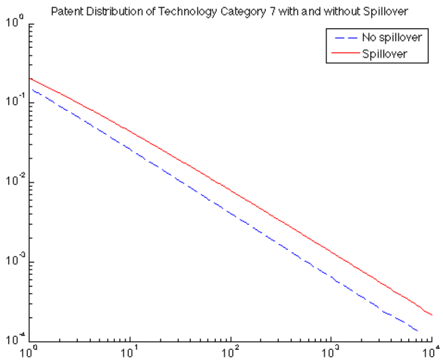
(2) Sector 2



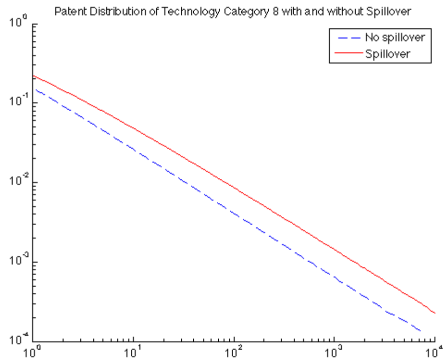
(3) Sector 4



(4) Sector 5

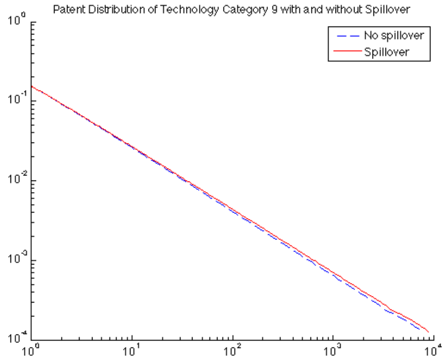


(5) Sector 7

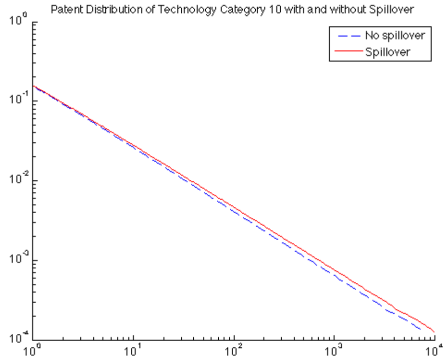


(6) Sector 8

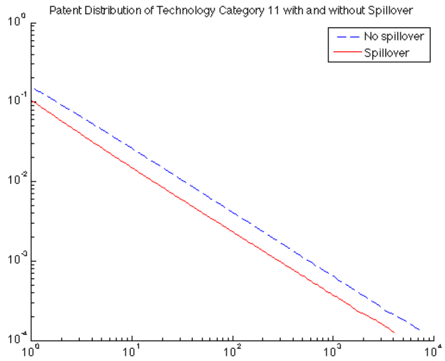
Figure A.2
Comparison of sectoral patent distribution with and without spillover effect acros sectors



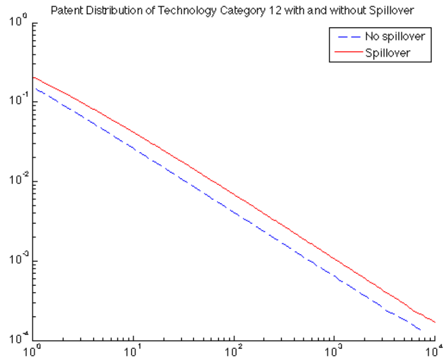
(7) Sector 9



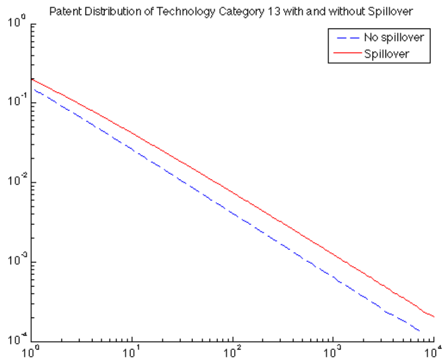
(8) Sector 10



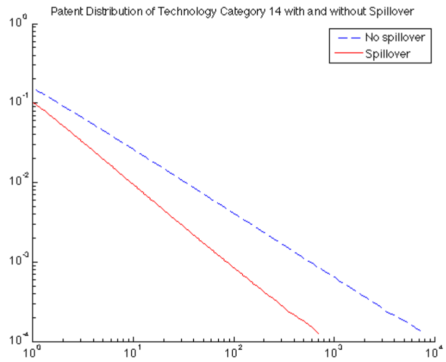
(9) Sector 11



(10) Sector 12

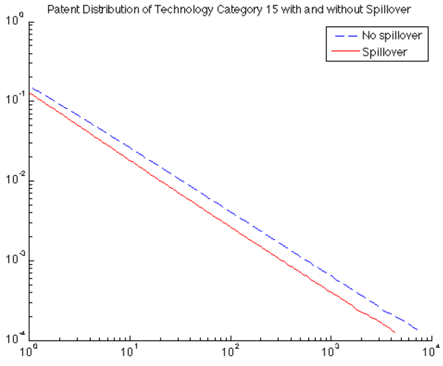


(11) Sector 13

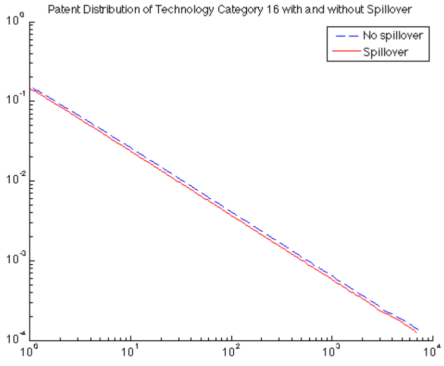


(12) Sector 14

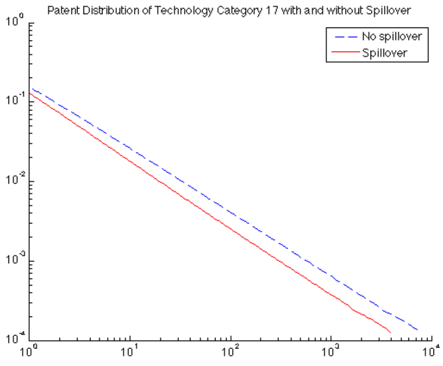
Figure A.2
Contined



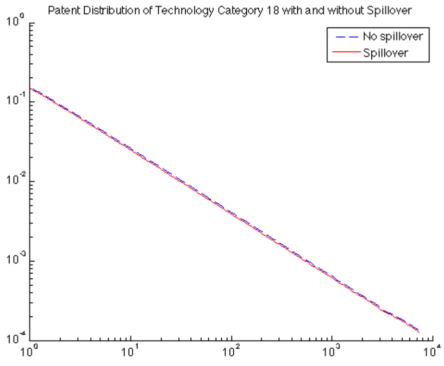
(13) Sector 15



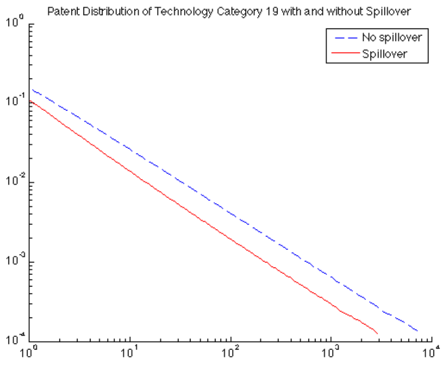
(14) Sector 16



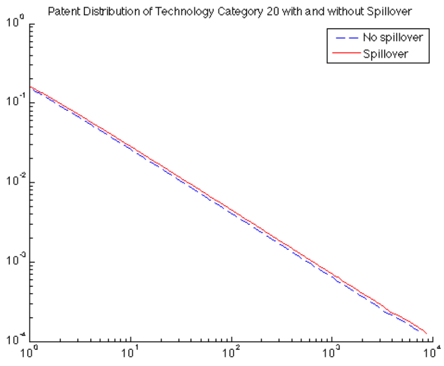
(15) Sector 17



(16) Sector 18

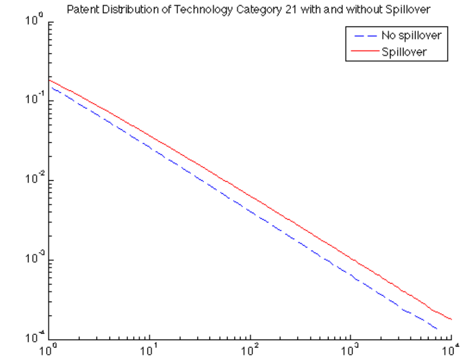


(17) Sector 19

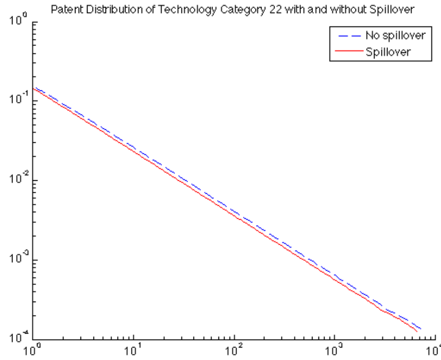


(18) Sector 20

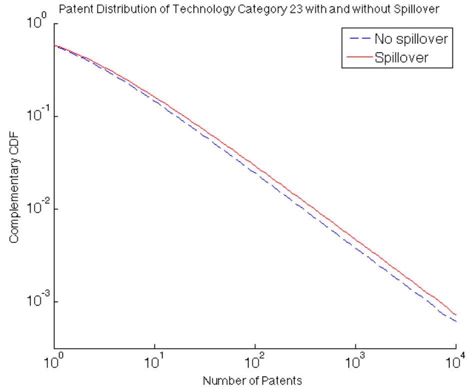
Figure A.2
Contined



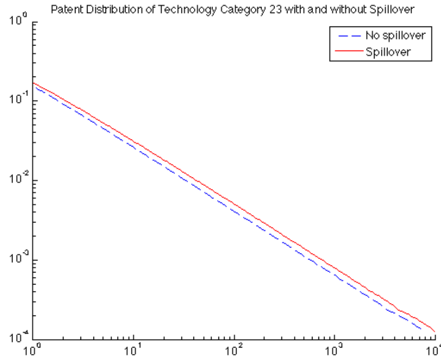
(19) Sector 21



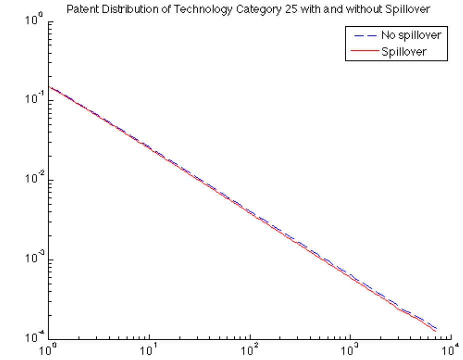
(20) Sector 22



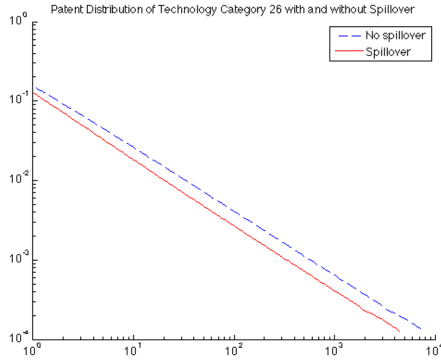
(21) Sector 23



(22) Sector 24

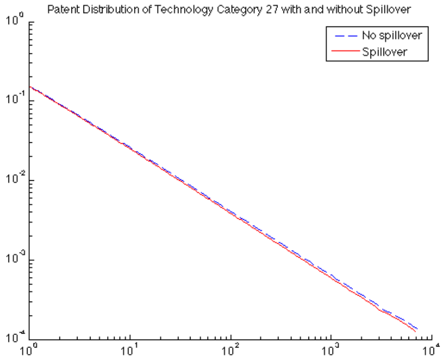


(23) Sector 25

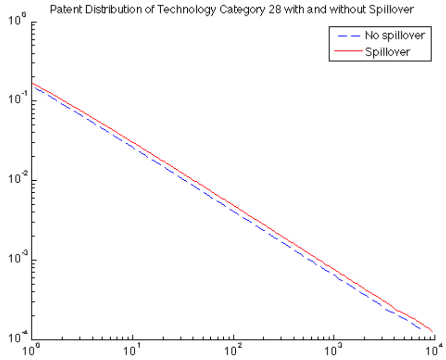


(24) Sector 26

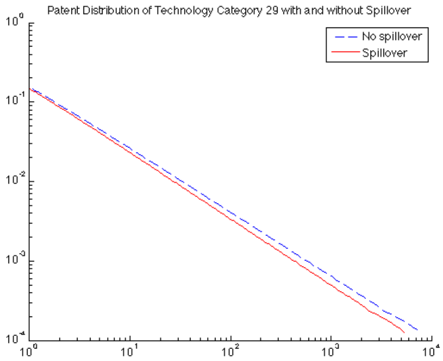
Figure A.2
Contined



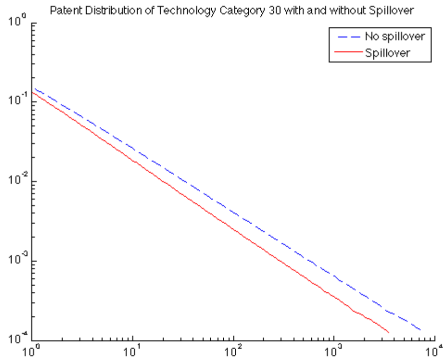
(25) Sector 27



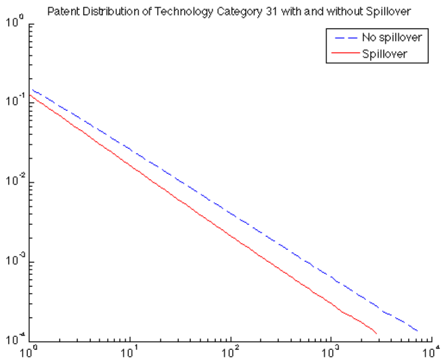
(26) Sector 28



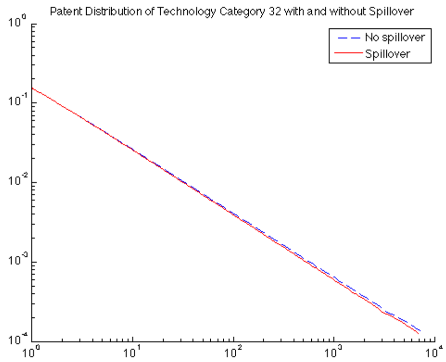
(27) Sector 29



(28) Sector 30

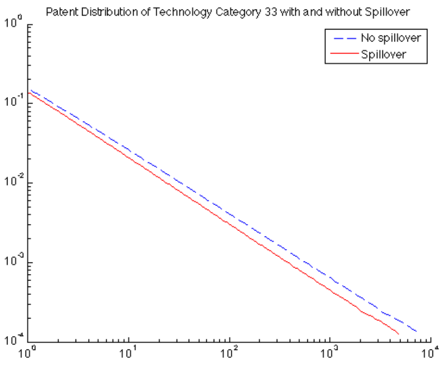


(29) Sector 31

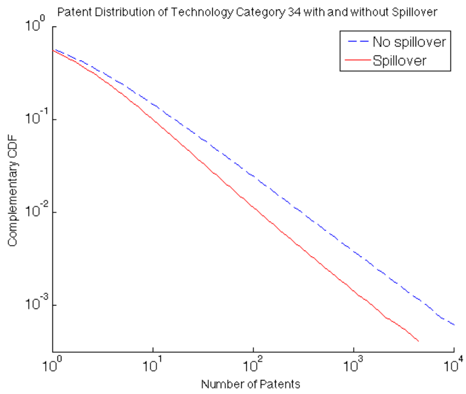


(30) Sector 32

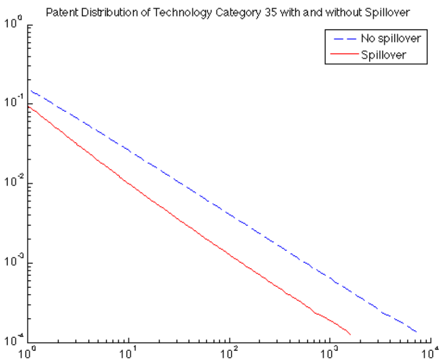
Figure A.2
Contined



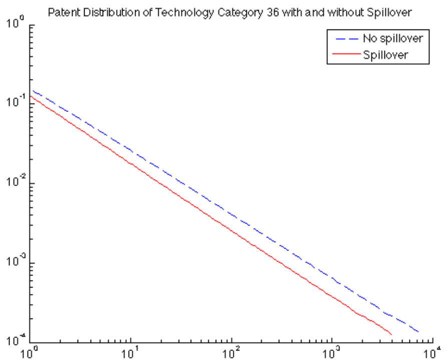
(31) Sector 33



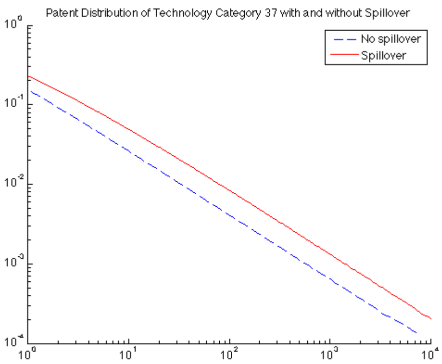
(32) Sector 34



(33) Sector 35

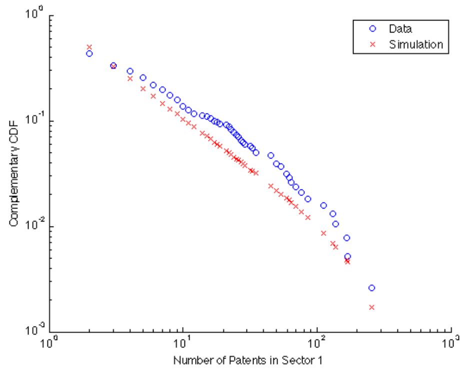


(34) Sector 36

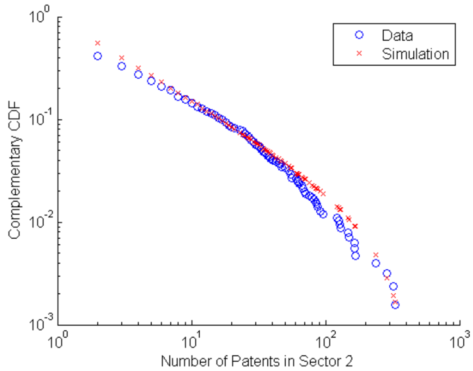


(35) Sector 37

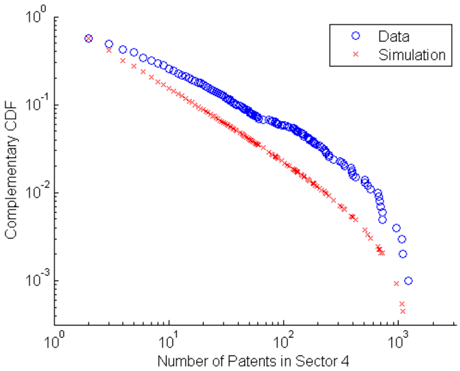
Figure A.2
Contined



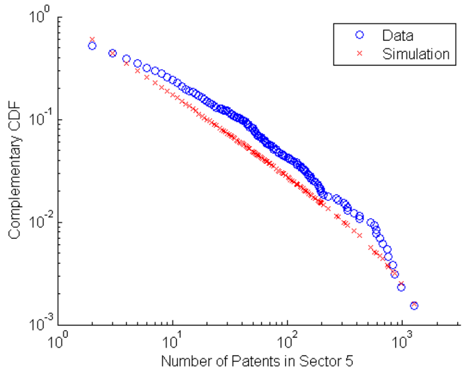
(1) Sector 1



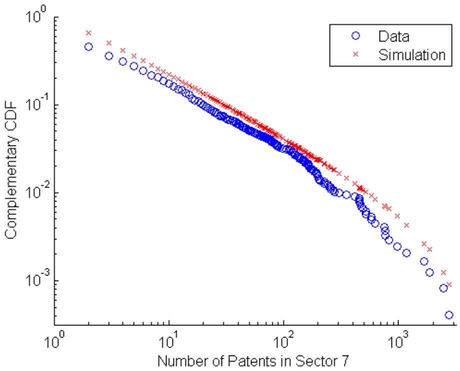
(2) Sector 2



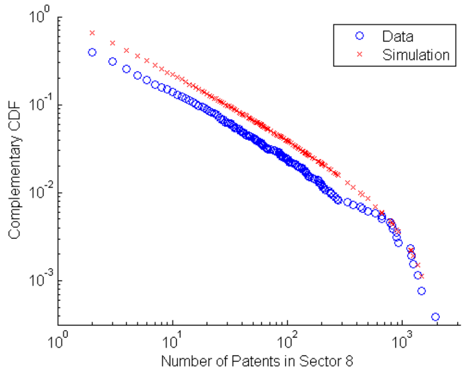
(3) Sector 4



(4) Sector 5

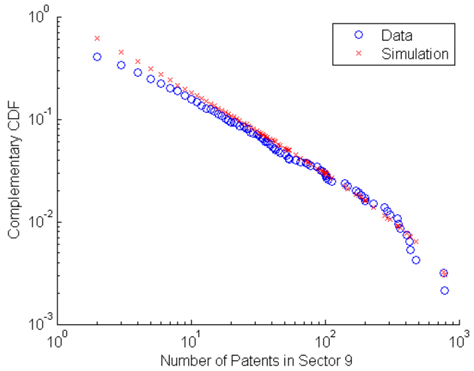


(5) Sector 7

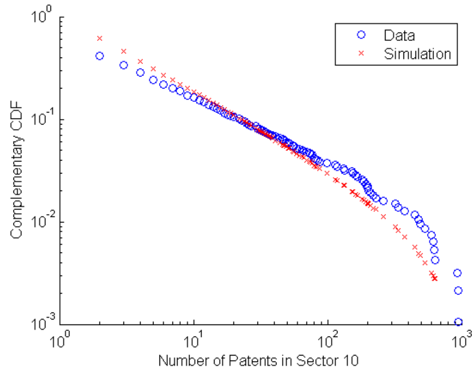


(6) Sector 8

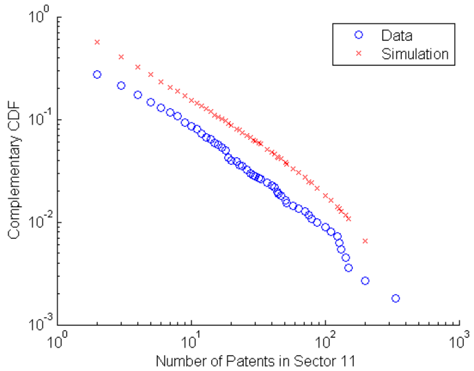
Figure A.3
PATENT DISTRIBUTION ACROSS SECTORS: SIMULATION V.S. DATA



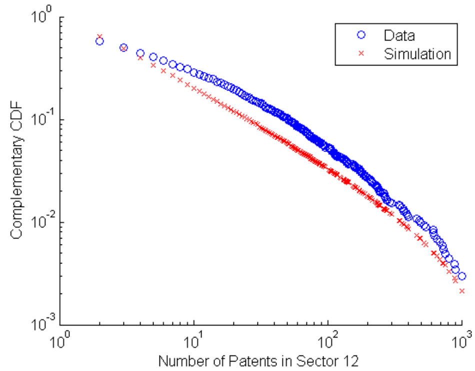
(7) Sector 9



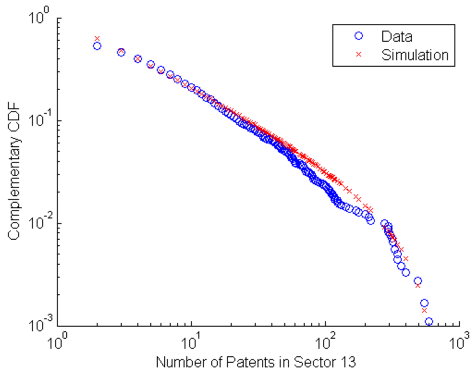
(8) Sector 10



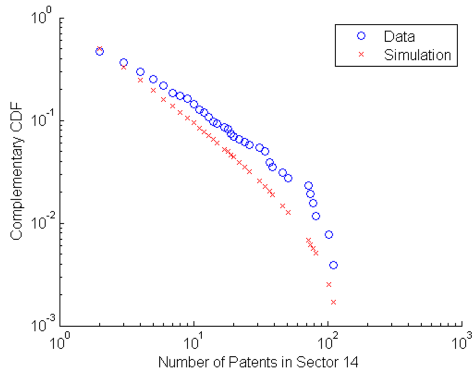
(9) Sector 11



(10) Sector 12

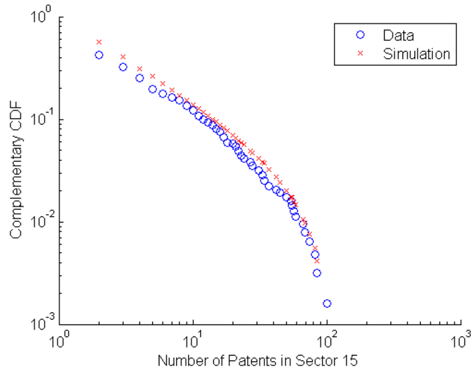


(11) Sector 13

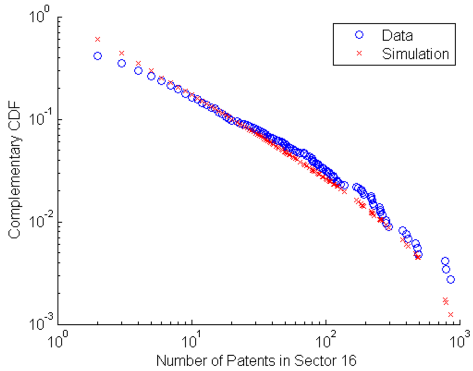


(12) Sector 14

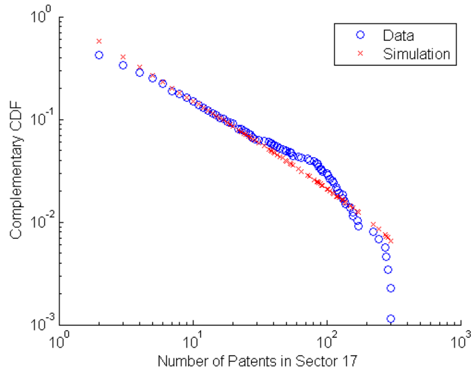
Figure A.3
Contined



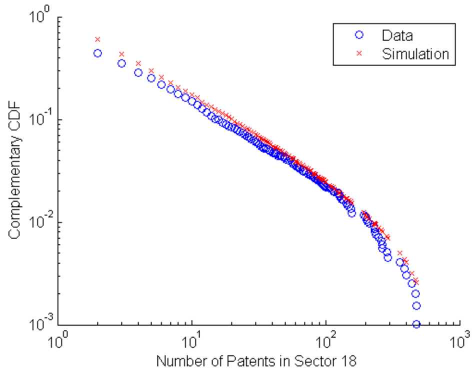
(13) Sector 15



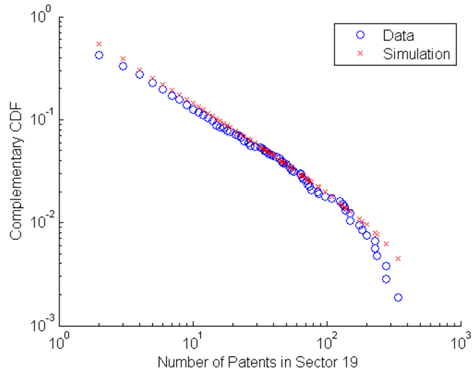
(14) Sector 16



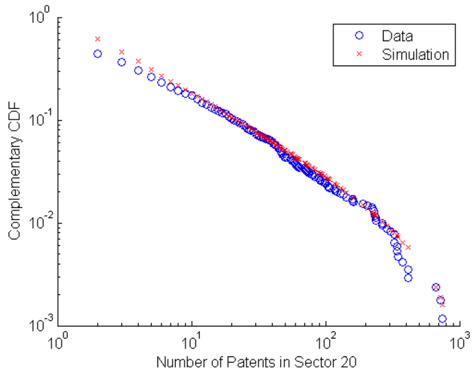
(15) Sector 17



(16) Sector 18

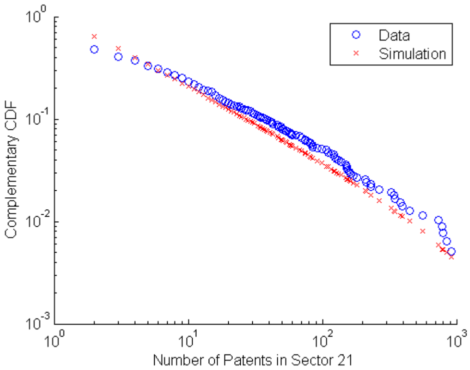


(17) Sector 19

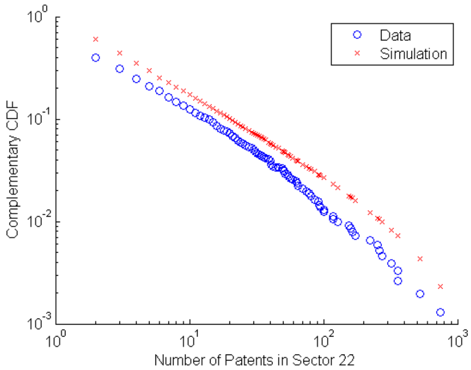


(18) Sector 20

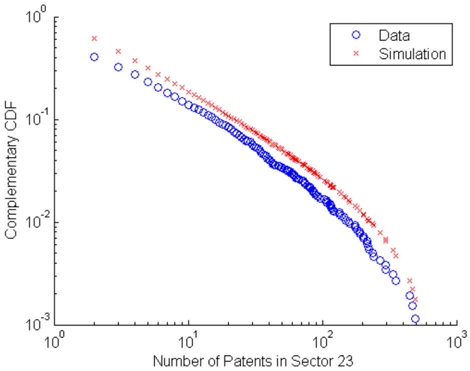
Figure A.3
Contined



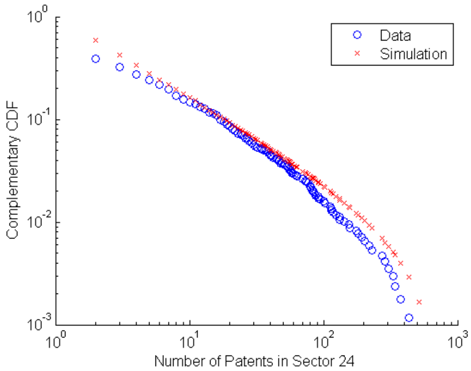
(19) Sector 21



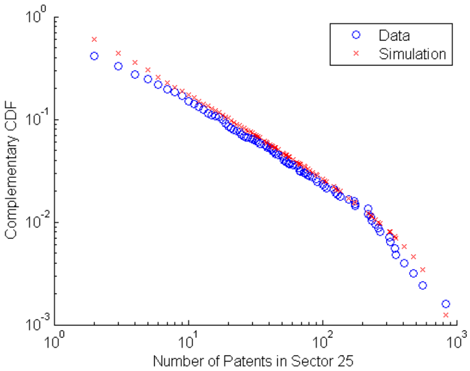
(20) Sector 22



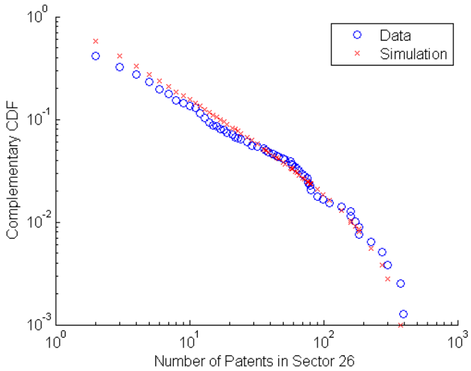
(21) Sector 23



(22) Sector 24

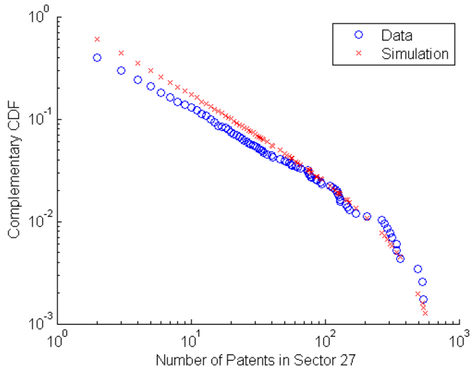


(23) Sector 25

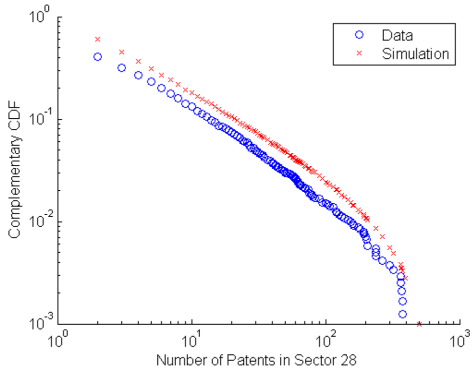


(24) Sector 26

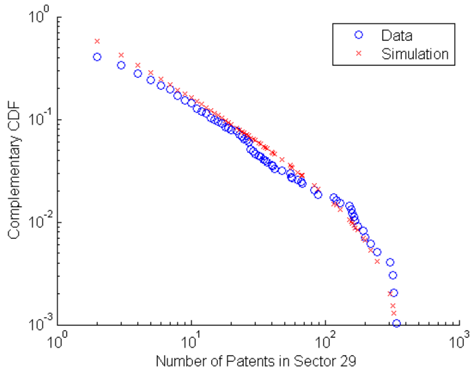
Figure A.3
Contined



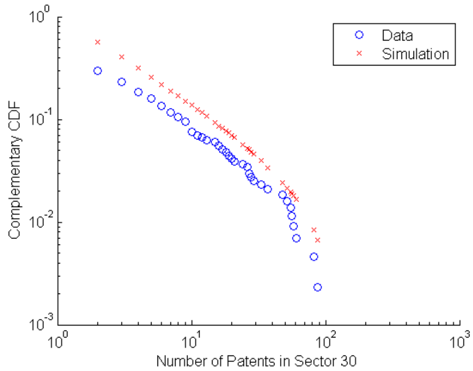
(25) Sector 27



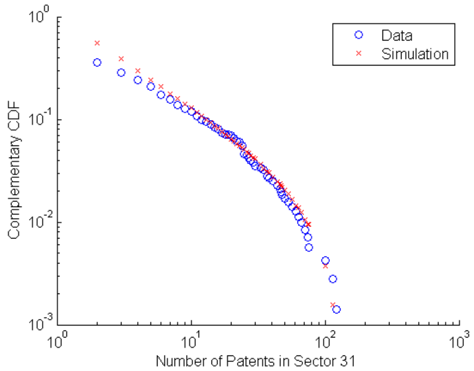
(26) Sector 28



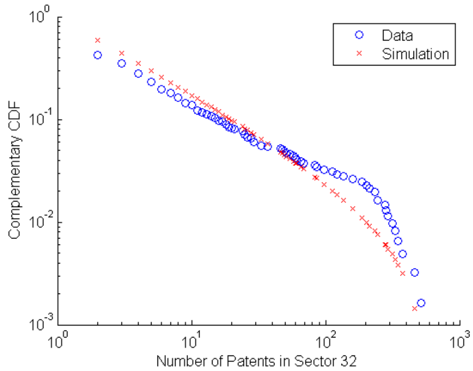
(27) Sector 29



(28) Sector 30

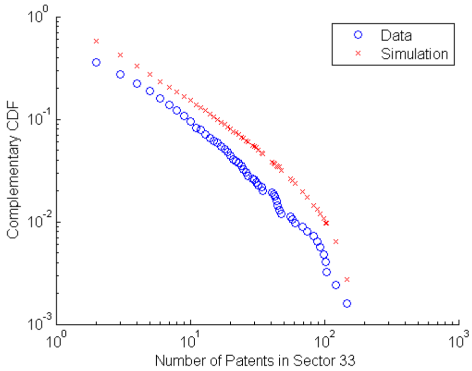


(29) Sector 31

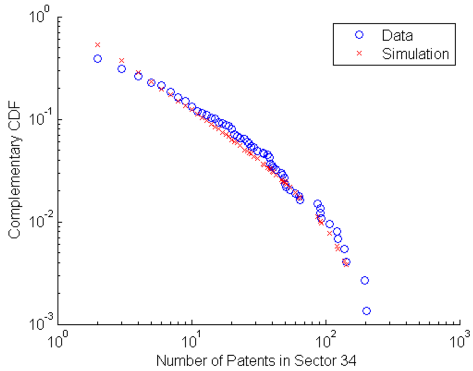


(30) Sector 32

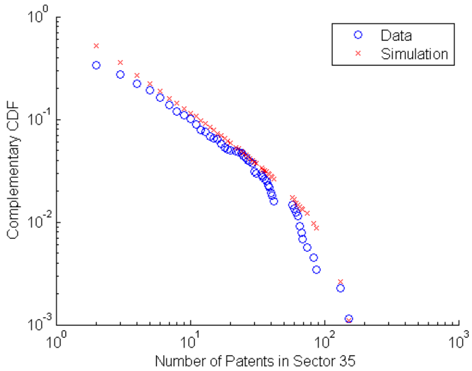
Figure A.3
Contined



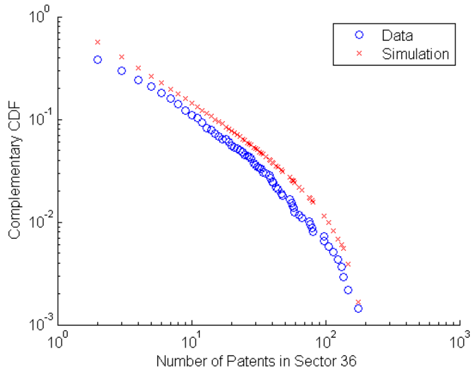
(31) Sector 33



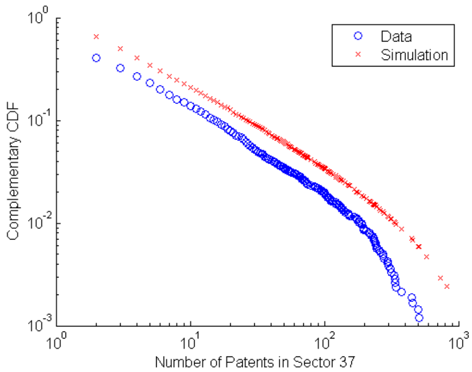
(32) Sector 34



(33) Sector 35



(34) Sector 36



(35) Sector 37

Figure A.3
Contined

References

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. (2010). Growth through heterogeneous innovations.

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*.

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

David, P. A. (1990). The dynamo and the computer: An historical perspective on the modern productivity paradox. *The American Economic Review*, (2):355.

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.

Newman, M. E. (2004). Analysis of weighted networks. *Physical Review E*, 70(5):056131.

Oberfield, E. (2011). 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.

Rosenberg, N. (1990). *Inside the black box : technology and economics.* Cambridge Univ. Press.

Rossi-Hansberg, E. and Wright, M. L. (2007). Establishment size dynamics in the aggregate economy. *The American Economic Review*, pages 1639–1666.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics*, (2):331–360.

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