

Growth Through Intersectoral Knowledge Linkages

Jie Cai and Nan Li*

October 2012

Abstract

The majority of innovations are developed by multi-technology (or multi-sector) firms. The knowledge needed to invent new products is more easily adapted from some sectors than from others. Here, we study this network of knowledge linkages between sectors and its impact on firm innovation and aggregate growth. We develop a general equilibrium model of multi-sector firm innovation in which intersectoral knowledge linkages determine a firm's self-selection into different sets of sectors and its R&D investment across sectors. It captures how firms evolve in the 'technology space', accounts for cross-sector differences in R&D intensity, and describes an aggregate model of technological change. Using simulations, we demonstrate that the model can match new observations concerning firms' multi-technology patenting behavior documented in this paper. The model also yields new insights into the effects of barriers to diversity on growth.

Keywords: Endogenous growth; R&D; Intersectoral knowledge spillovers; Firm innovation; Multiple sectors; Resource allocation

JEL Classification: O30, O31, O33, O40, O41

*Cai: University of New South Wales, address: Department of Economics, University of New South Wales, email: april.cai@unsw.edu.au. Li, Ohio State University and International Monetary Fund, address: 700 19th Street NW, Washington DC 2043, email: li.854@osu.edu. Acknowledgement: We thank Ufuk Akcigit, Paul Beaudry, Bill Dupor, Chris Edmond, Sotirios Georganas, Joe Kaboski, Aubhik Khan, Sam Kortum, Amartya Lahiri, Roberto Samiengo, Mark Wright, as well as seminar participants at Brown University, London School of Economics, University of British Columbia, University of Melbourne, Ohio State University, the UNSW-University of Sydney joint macro seminar, NBER EFJK group meeting, Econometric Society Winter Meeting, ASSA-2012, Reserve Bank of Australia, Chinese University of Hong Kong and the IMF institute, for many helpful discussions and comments. The views expressed in this paper do not reflect those of the International Monetary Fund.

1 Introduction

Innovation hardly ever takes place in isolation. Technologies depend upon one another, yet vary substantially in their degree of applicability. Some innovations, such as the electric motor, create applicable knowledge that can be easily adapted to design new products in a vast range of sectors; while other inventions introduce technologies that are limited in their scope of application. The *interconnections* between different technologies and the stark contrasts in the way new technologies affect future innovation have long been recognized by economic historians.¹ The majority of theoretical works on endogenous growth, however, tend to treat innovations in different technologies as isolated from each other and equally influential.²

Empirical evidence presented in this paper suggests that intersectoral knowledge spillovers are *heterogeneous* and highly *skewed*: a small number of technology categories are responsible for fostering a disproportionately large number of subsequent innovations in the economy. Meanwhile, a close inspection of the firm patenting data points to the importance of innovations by *multi-technology firms* which are able to internalize intersectoral spillovers: More than 42% of patenting firms innovate in more than one technological area, accounting for 96% of innovations in the economy. And among all patent citations, 52% are made between distinct technologies, indicating substantial knowledge spillovers across sectors.³

The questions are: How do firms decide on what kinds of technologies to develop, in which sectors to apply their existing technologies and grow their business? How does the technological progress in one sector transmit to another? And ultimately, what are the *aggregate implications for growth* of the technological diversification of firms? The consequences for growth of government policy directed at stimulating innovations in certain sectors hinges on better understanding of the above questions. Addressing these questions requires a structural framework that integrates micro empirical observations into a macro-growth model with multiple sectors and heterogeneous firms (in terms of technology scope).

This paper therefore endeavors to achieve two goals. The first goal is to develop a general

¹David (1991), Rosenberg (1982), Landes (1969), for example, emphasize the dramatic growth impact played by general-purpose technologies (GPTs).

²Notable exceptions can be found in the literature of GPTs (see Jovanovic and Rousseau, 2005; Helpman, 1998, and Bresnahan and Trajtenberg, 1995). However, our paper focuses on understanding the impact of technology linkages on firm innovation and growth, where the notion of applicable technologies is related to but distinct from the concept of GPTs.

³This is based on 428 technology classes (U.S. Patent Classification System) provided by U.S. Patent and Trade Office for the period 1976-2006. The percentage becomes even higher when using more disaggregated classifications. Similarly to this observation, using the product-level data, Bernard, Redding and Schott (2011) find that 41% of U.S. manufacturing firms operate multiple product lines, accounting for 91% of total sales. Recent trade literature has also highlighted that world trade flows are dominated by multi-product firms (e.g. Mayer, Melitz and Ottaviano, 2011). Therefore, understanding how firms expand their technology and product range sheds light on both technological progress and aggregate production and trade, although this paper focuses on the former.

equilibrium model of multi-sector (or multi-technology) firm innovation to address these issues.⁴ The framework is built on the leading models of endogenous growth which unify firm-level studies of R&D, patenting and firm growth with aggregate analysis of technological change, originally developed by Klette and Kortum (2004). Relative to the existing work in this literature, our framework emphasizes two new features: *heterogenous intersectoral knowledge linkages*⁵ which determine the productivity of R&D when adapting knowledge from one sector to another and consequently affect firms' R&D distribution over multiple sectors; and the *barriers to diversification* (sectoral fixed costs) which cause sequential entry and exit of firms in different sectors. The model captures how firms evolve in the 'technology space', and describes how knowledge accumulates in different sectors and in the aggregate economy. It yields simple expressions relating growth to cross-sector knowledge circulation in the economy. In particular, it points out that barriers to diversity prevent firms from fully internalizing cross-sector spillovers, reducing technological progress.

The second objective is to assess the performance of the model using patent and R&D data. Technology interconnections are conceptual and difficult to measure. We propose several measures based on the patent citation network linking the knowledge receiving and sending sectors. The method establishes a particular hierarchy in technology space that is amenable to empirical explorations. Based on these measures and using firm patenting data, we document several new observations concerning firms' multi-technology patenting: (i) firms with more patents or innovating in more patent classes are more concentrated in highly applicable technologies, (ii) yet, their recent patents are in the less applicable ones, and (iii) firms whose initial technologies are more applicable innovate faster. When simulated using a large panel of firms innovating in heterogenous (multiple) sectors, our model reproduces each of the new facts above, as well as endogenously generates Pareto firm size distribution, in line with existing empirical findings.

In the model, firms invent new products by conducting R&D to adapt existing knowledge in various technologies. Applicable technologies enhance the innovational productivity of R&D in downstream sectors and contribute to a sequence of innovations in various areas—exhibiting 'innovational complementarities'. Therefore, when firms choose R&D optimally, the model predicts that the equilibrium value associated with the introduction of a new innovation is not confined to

⁴The terms technologies, technological categories and sectors are interchangeable throughout the paper. One sector embodies one specific type of technology. Although distinguishing a firm's position in technology space and product space is interesting for certain issues previously explored in Bloom et al. (2010), it is not the interest of this paper.

⁵We focus on the 'deep' knowledge linkages between technologies which are due to intrinsic characteristics of technologies, and do not vary over time. For this reason, we summarize citations made to (and from) patents that belong to the same technology class over 30 years to form the intersectoral knowledge diffusion network. We also abstract from other differences across sectors, such as demand and production costs in the model. While assuming so may limit its fit, it is in line with Nelson and Winter (1977), who argue that innovations follow 'natural trajectories' that have a technological or scientific rationale rather than being fine tuned to changes in demand and cost conditions.

its own future profit gains; rather, it also depends on the *application value* of this new technology in *all* sectors.

For any given sector in which a firm intends to enter or continue conducting research, a period-by-period firm idiosyncratic fixed cost is required. These fixed costs, acting as barriers to diversity, make research in multiple sectors a *self-selection* process: a firm develops new products in sectors where it can most efficiently utilize its existing range of knowledge portfolio. Consistently with the evidence in Section 2, firms conducting research in multiple areas are more likely to be concentrated in highly applicable technologies, because they are better at internalizing intersectoral knowledge spillovers and have stronger incentives to invest in these areas.

Although the sectors with high knowledge application value attract firms to invest intensively in R&D, the model suggests that a counteracting force is at play: namely, the fierce competition in these sectors. Firms would only conduct research and operate in a sector if the expected application value of its technology is large enough to cover the fixed cost of research.⁶ Therefore, smaller firms with less knowledge capital start from sectors with high application value—what we call ‘central sectors’ in the technology space—whereas firms with larger knowledge capital in multiple sectors expand into technologies with lower applicability but allowing them to have larger market shares—what we call ‘peripheral sectors’. The tradeoff between *innovational applicability* and *product market competition*—which is at the heart of the R&D resource allocation mechanism in the economy—leads to a stationary firm distribution across sectors and a stationary (normalized) sector size distribution on the balanced growth path.

Innovation by its nature can be highly uncertain. In the model we assume that firms face idiosyncratic shocks to the success rates of innovations and the fixed costs of research in various sectors. Therefore, even though firms on average enter multiple sectors sequentially—that is, firms typically start from central sectors and slowly venture into periphery after accumulating enough private knowledge in related sectors—not all firms follow the same sequence. In any given sector, existing firms innovate, expanding their sizes as they create new varieties, and exit after experiencing a sequence of negative innovation shocks or high fixed costs. In addition, new firms enter if they have accumulated enough knowledge capital in related sectors. This process *endogenously* generates a distribution of firm size in each sector (and in the whole economy), converging to a Pareto distribution in the upper tail, in line with existing empirical findings of firm size distribution.⁷

⁶Firms in the model are subject to idiosyncratic sectoral innovation shocks, which is i.i.d over time and across firms. Thus, a firm exits a specific sector if it experiences a range of negative shocks such that the expected payoff of operating in that sector cannot cover the fixed cost. As will be shown in Section 4.1, these idiosyncratic shocks also help to ensure a stationary Pareto firm size distribution.

⁷Firm or establishment-level data shows that firm size distributions within narrowly defined sectors and within the overall economy are widely dispersed and follow a Pareto distribution, as documented in Axtell (2001), Rossi-Hansberg and Wright (2007) and Luttmer (2007).

Because knowledge in different sectors is related (by various degrees), firms can expand through the technology space by developing new knowledge close to their existing technology mix. When the scope and applicability of a firm’s knowledge increases, so do the opportunities to innovate, profit and grow in related sectors. Moreover, existing sectors also benefit from the growth in the new innovating sectors as a consequence of knowledge spillovers in the opposite direction. With controls for the size of knowledge capital, the model predicts that firms concentrating more on applicable technologies tend to innovate faster. Again, this is consistent with the firm-level observations documented in Section 2.

In addition to explaining firm-level observation, our model predicts that equilibrium R&D intensity is higher in sectors with larger value of knowledge capital. While this value is not directly observable, the model suggests that it increases with its innovational applicability—the ability to foster technical advances in a wide variety of sectors. Employing U.S. Compustat firm R&D data, we find that R&D intensity in sectors with highly applicable knowledge is, indeed, larger.

The model also yields new insights regarding the mechanism through which barriers to diversity reduce technological progress in the multi-sector environment.⁸ In the presence of these barriers, only a small number of firms can afford a sequence of fixed costs, and innovate in multiple sectors. Thus, these barriers directly block the knowledge circulation in the entire technology space by preventing firms from fully internalizing spillovers from other sectors, imposing a negative effect on aggregate technical advances.

Related Literature Our paper is most closely related to Klette and Kortum (2004), which connects theories of aggregate technological change with findings from firm and industry-level studies of innovation. However, the growth implications of technology diversification in the presence of interconnections have been largely unexplored.⁹ In the past, most theoretical works on endogenous growth (e.g. Romer, 1986,1990; Lucas, 1988; Segerstrom, Anant and Dinopoulos, 1990; Aghion and Howitt, 1992; Grossman and Helpman, 1991a, 1991b; and Jones, 1995) and research on in-

⁸The economic channel through which entry costs decrease growth in this paper is very different from the ones stressed in the previous literature. For example, entry costs discourage small but innovative entrepreneurs from entry and keep existing establishment inefficiently large in Boedo and Mukoyama (2009), or they distort the allocation of talent across sectors as in Buera, Kaboski and Shin (2011). In Barseghyan and DiCecio (2010), higher entry costs reduce productivity of the marginal entrant through a general equilibrium effect on wage.

⁹The previous working paper version of Ngai and Samaniego (2011) suggests including cross-industry spillovers into their model which identifies factors that account for endogenous differences in research activities and productivity growth across sectors. However, they abstract from this because they find that citations are dominated by within-industry citations, speculating that cross-industry spillovers are small. In this paper we show that, although small in absolute quantity, the spillovers are large enough to have significant implications on growth through higher-order linkages. Furthermore, previous empirical studies also suggest so. For example, using firm-level R&D investment data in five high-tech industries, Bernstein and Nadiri (1988) find that knowledge spillovers largely vary across sectors and are highly significant. Wieser (2005), in his survey paper, claims that spillovers between sectors are more important than those within sectors, when considering both the social and private return of R&D.

novation and firm dynamics (e.g., Klette and Kortum, 2004; Luttmer, 2007, 2012 and Atkeson and Burstein, 2010) have considered a single type of technological change or implicitly assumed a perfectly homogeneous technology space in the sense that innovation takes place in any sector with equal probability. There are no explicit interactions between different sectors or distinctions between technologies with different degrees of applicability, and hence, no room to discuss a firm’s technological position and its impact on future growth.¹⁰

Empirical work by Jaffe (1986), on the other hand, suggests that firms’ technology position provides different technological opportunities that matter for firms’ innovative success. In that paper, however, technological position is taken as exogenous and can only be changed over long time periods. Our study advances Jaffe’s findings by constructing a structural model which allows for the *endogenous sorting* of firms across technologies, providing further understanding of the relationship between technological linkages and firms’ dynamic decisions in allocating their research effort. Using firm-product-level observations, Bernard, Redding and Schott (2009, 2010) finds that most firms switch their product frequently, and that endogenous product selection has important quantitative implications on measured firm and aggregate productivity. Obviously, our focus is entirely different: we examine firm innovation behavior instead of production performance. The more interesting difference is that we allow sectors to be inherently connected by their knowledge spillovers. Hence, sector selection also depends on the firm’s existing position in the technology space.

Distinguishing between different types of research and their impact is currently being pursued in a number of papers. Distinguishing between basic research and applied research, Akcigit, Hanley and Serrano-Velarde (2011) focuses on analyzing the impact of the appropriability problem on firms’ incentives to conduct basic research.¹¹ We do not limit our analysis to two distinct types of research; rather, we consider the richer and more complex structures of technological interdependence across *multiple* technologies and also quantify the strength of knowledge linkages using cross-sector citation data.¹² Our purpose is to integrate firm multi-technology innovations and the cross-sector spillovers into the endogenous growth models. Akcigit and Kerr (2010) studies how exploration versus exploitation innovations affect growth. Akin to this notion, Acemoglu and Cao (2010) considers incremental R&D engaged in by incumbents and radical R&D undertaken by potential entrants. Similarly to Acemoglu and Cao (2010), our paper also allows for simultaneous

¹⁰For example, in the expanding variety models (Romer, 1988 and Grossman and Helpman, 1991a), the initial varieties do not affect the expected productivity in producing or R&D in another sector. In Aghion-Howitt (1992) quality ladder model quality improvement takes place across all products at the same time.

¹¹One of the empirical facts documented is related to this paper: firms’ multi-industry presence is positively associated with their devotion to basic research.

¹²Although cross-sector knowledge spillover is possible in their model, the magnitude of spillovers across different sectors is homogeneous.

innovations by continuing firms and entrants; however, the different technological fields in which large versus small firms (or incumbent versus entrants) choose to innovate in our model reflect an endogenous equilibrium outcome.

Our work also builds on earlier literature in development economics that emphasizes the role of sectoral linkages and complementarity in explaining growth (see Leontief, 1936 and Hirschman, 1958). Previous work in this area typically focuses on vertical input-output relationships in production between sectors—as in Jones (2011), and export-based measures of product relatedness—as in Hidalgo, Klinger and Hausmann (2007) and Hausmann, Hwang and Rodrik (2007).¹³ This paper focuses on linkages dictated by their knowledge content, which is more suitable for understanding the mechanics of technological innovation.

Finally, this paper also adds to previous works studying the determinants of cross-industry differences in R&D intensity. Klenow (1996) evaluates the implications of three hypotheses—technological opportunity, market size and appropriability in an extended model of Romer (1990). Ngai and Samaniego (2009), based on a calibrated model, finds that differences in R&D intensity mainly reflect technological opportunities (interpreted as the parameter of knowledge production governing decreasing returns to research activity). Empirical evidence and the model developed in this paper both suggest that these differences can be attributed to sectoral technology applicability—which constitutes a direct measure and interpretation of technological opportunity, complementing previous findings.¹⁴

The paper begins by presenting some new sector and firm-level findings to motivate our modeling approach. The model itself is developed in Section 3. We then describe the aggregate properties of the stationary balance growth path equilibrium in Section 4. Section 5 discusses calibration and parameterization of the model, and the results from simulations. Section 6 discusses possible directions of future research and policy implications.

2 Empirical Underpinning

In this section, we first document several empirical observations that motivate our model using patent citations, firm patenting and R&D investment data. First, we show that the applicability of different technologies is heterogenous and highly skewed. The applicability measure we constructed for different sectors is found to be positively correlated with the sectoral R&D intensity (defined

¹³Other research studies the role of input-output relationship in understanding sectoral co-movements and the transmission of shocks over the business cycle, such as Lucas (1981), Basu (1995), Horvath (1998), Conley and Dapor (2003), Carvalho (2010) and recently, Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012).

¹⁴Other contributions in this literature include Pakes and Schankerman (1984), Levin et al (1985) and Jaffe (1986, 1988).

as R&D expenditure over sales). Next, using the measure of technology applicability, we document several facts concerning firms' multi-patenting behavior and the relationship between firms' technological position and innovation performance.

Our main data source is the 2006 edition U.S. Patent and Trade Office (USPTO) data from 1976 to 2006.¹⁵ We focus on firm patenting activities in this paper, as the model is designed to mainly understand firm innovation behavior.¹⁶ Patent applications serve as proxies of firms' innovative output, and their citations are used to trace the direction and intensity of knowledge flows within and across technological classes.¹⁷ Each patent corresponds to one of the 428 3-digit United States Patent Classification System (USPCS) technological field (NClass). Another source of data is from U.S. Compustat 1970-2000 which includes firm-level R&D expenditure and firm performance data. We use this information to construct sector-level R&D intensity. Additional information about the data and construction of various measures appears in Appendix A.

2.1 The Measurement of Technology Applicability

Network of Intersectoral Knowledge Linkages We summarize cross-sector citation flows connecting different technology classes to form the intersectoral knowledge diffusion network.¹⁸

The following example illustrates the highly heterogeneous nature of technology interconnections. Consider a network depicted in Figure I consisting of eight interacting technologies. Every vertex corresponds to one type of technology, and every arrow indicates the direction of the knowledge flow. In this example, knowledge created in 1 can be adapted to develop new knowledge in 1-7, indicating high technology applicability. In contrast, knowledge in 8 and 9, for example, are sector-specific and cannot be applied to any others. 2 is a more important knowledge contributor than 6. Even though 2 contributes to fewer sectors; the one sector to which its knowledge is applicable is an important sector (sector 1) and hence is of high value. Similarly, 3 should be ranked higher than 4 as it influences 1 *indirectly* through 2.

The actual network of intersectoral knowledge linkages (shown in Figure II) based on citations between 428 technology classes resembles this network structure. It exhibits strong heterogeneity: a small fraction of sectors play a disproportionately important role in fostering subsequent innovations

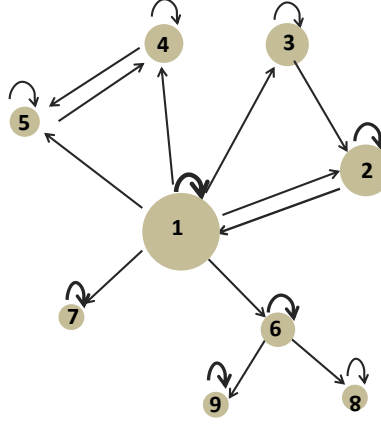
¹⁵See Hall, Jaffe and Trajtenberg (2001) for detailed description of the data.

¹⁶Merging firm patent data and U.S. Census firm-level data, Balasubramanian and Sivadasan (2011) finds that although only 5.5% of all manufacturing firms engage in patenting activity, they play an important role in the aggregate production, accounting for about 60% of value added. Therefore, understanding the behavior of patenting firms substantially improves our understanding of the driving force of growth.

¹⁷We only consider patents by domestic and foreign non-government institutions.

¹⁸Since we are interested in studying the 'deep', time-invariant characteristics between different technologies, which firms take as exogenously given, we adopt patent citation data spanning the 1976-2006 period to form the network. Pooled citations of 30 years also help to even out noises in the annual citation data. At the US 3-digit patent classification level, one-third technology pairs never cite each other, implying that many technologies are truly unrelated.

Figure I: The Intersectoral Network Representations of Knowledge Linkages: An Example

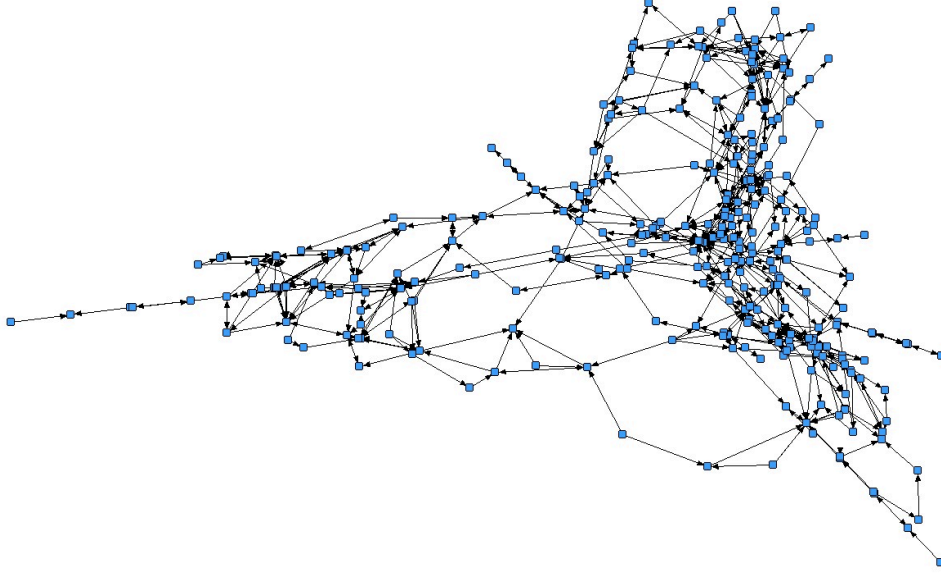


in other sectors.

Calculating Sector-Specific Technology Applicability The relationships of knowledge complementarity (especially, the higher-order interconnections) make it difficult to evaluate the contribution of any single innovation to the whole technology space. Hence, the first challenge is to construct such a sector-level measure that characterizes the importance of different sectors as knowledge suppliers to their immediately application sectors as well as their role as indirect contributor to chains of downstream sectors. Therefore, the citation count (or variation of it such as Garfield’s (1972) “impact factor” or the forward-citation-weighted count) may be a poor proxy of what is really of interest.

To handle this issue, we apply Kleinberg’s (1998) iterative algorithm—which is proved to be the most efficient at extracting information from a highly linked environment—to the knowledge diffusion network. We construct a measure quantifying the *applicability* of each technology (denoted by a^i in sector i)—called ‘authority weight’ in Kleinberg’s original work. This algorithm is a fixed-point iteration which generates two inter-dependent indices for each node in the network: authority weight (aw^i)—the ability of contributing knowledge to the entire network; and hub weight (hw^i)—the ability of absorbing knowledge. Let J be a set of technological categories. A citation matrix for J is a $|J| \times |J|$ nonnegative matrix $(c^{ij})_{(i,j) \in J \times J}$. For each $i, j \in J$, c^{ij} denotes the number of citations to sector i made by j and $S = (S^i)_{i \in J}$ represents the vector of patent stock in different

Figure II: Intersectoral Network Corresponding to Patent Citation Share Matrix



Notes: NBER patent citation data, 428 technological categories (NClasses). A (directed) link is drawn for every citation link that counts more than 5% of the total citations made by the citing sector.

sectors. Formally, they are calculated according to:

$$aw^i = \lambda \sum_{j \in J} W^{ij} hw^j$$

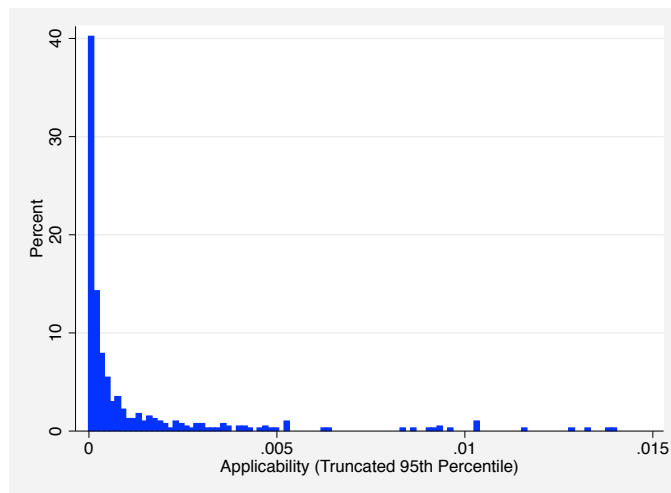
$$hw^i = \mu \sum_{j \in J} W^{ji} aw^j$$

where λ and μ are the inverse of the norms of vectors $(aw^i)_{i \in J}$ and $(hw^i)_{i \in J}$, respectively. W^{ij} denotes the weight of the link, corresponding to the strength of knowledge contribution by sector i to sector j . As noted in Hall et al. (2001), sectors vary in their propensity to patent. We consider two ways to calculate the weight W^{ij} . First, $W^{ij} = 1$ if there is at least one citation made by j to i and zero otherwise. Thus, the weight is independent of the relative size between i and j . The second measure normalizes cross-citation counts by the number of patents in the citing class: $W^{ij} = c^{ij}/S^j$, reflecting the average rate at which patents in class j cite patents in class i .¹⁹ Most of the results reported below are based on the first method, except when information is only available at the 42 industrial sector level. At this level of aggregation, cross-citation exists for most sector-pairs and hence difficult to rank the importance of sectors. In this case, we use the second method.

¹⁹This method of constructing the weight is similar to the construction of production input-output matrix by calculating the share of input used by a downstream sector from any given upstream sector.

For robustness, we also construct three alternative measures to rank sectors’ technology applicability: the number of inward citations from other technology classes, the (weighted) average shortest distance to other technologies in the network, and the ‘upstreamness’ of a technology in the network²⁰ (See Appendix A for details of the construction of these measures). It turns out our four measures of technology applicability are highly correlated with each other and the results in the following sections are robust to using alternative measures.

Figure III: Distribution of Technology Applicabilities (aw) Across Sectors



Notes: The applicability measure is constructed by applying Kleinberg’s algorithm to the cross-sector patent citation network (NBER Patent Dataset, 1976-2002).

Distribution of Technology Applicability The first fact is that technology applicability is highly heterogenous across sectors. Figure III shows the highly skewed distribution of our measure of applicability across sectors, with a small number of sectors acting as the knowledge ‘authority’ in the technology space.

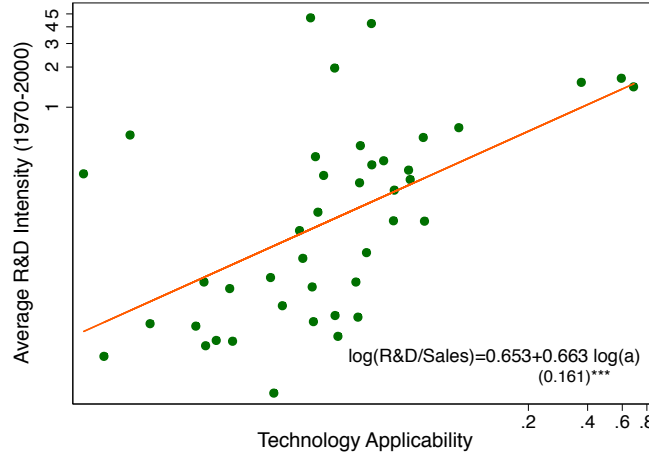
Technology Applicability and R&D Intensity It has been documented previously in the literature that there are large and persistent cross-sector differences in R&D intensities. Using a calibrated model, Ngai and Samaniego (2011) argue that these differences mainly reflect different ‘technological opportunities’.²¹ Here, we empirically investigate this relationship using our measure of technology applicability embodied in different sectors, giving a natural interpretation of the technological opportunity.

²⁰The ‘upstreamness’ measure is constructed using a similar method as in Antras et al (2012), modified to fit the patent citation data. The idea is that a more upstream technology class acts as a more intensive knowledge supplier to innovations in other categories.

²¹They capture technological opportunities as the parameters of knowledge production in their theoretical model.

Following the literature, we measure the long-run sectoral R&D intensity as the median ratio of R&D expenditures to sales among firms in Compustat dataset over the period 1970-2000.²² Besides the detailed technology categories, USPTO dataset also includes patent categories at the 2-digit SIC level. We use this information to assign patents to industry classes, construct our measure of technology applicability at the 2-digit SIC level and link this measure to R&D intensity in the same sector.

Figure IV: Sectoral R&D Intensity Significantly Increases with Its Technology Applicability



Notes: R&D intensity is the median ratio of R&D expenditures to sales among firms in Compustat over the period 1970-2000. The applicability measure is constructed based on cross-sector patent citation network (NBER Patent Dataset, 1976-2002). Both horizontal and vertical axes are in log scale. The solid line represents the fitted values. The brackets under the regression coefficient estimates shows the standard errors for the estimates.

Figure IV shows there is a strong positive relationship between the R&D intensity in a sector and the applicability of its technology. In later sections of the paper, we show that, in our model, firm optimal R&D decision leads to positive correlation between the sectoral R&D intensity and the application value of the knowledge embodied in the sector, thus explaining this observation.

2.2 Firm's Technological Position and Innovation Performance

Each firm is identified by its overall patent stock (S_f), and its technological position (P_f) which is independent of firm scale. Following Jaffe (1986), we characterize the firm's technological position by the *distribution* of its patents over all patent classes, defining a vector $P_f = (P_f^1, P_f^2, \dots, P_f^{428})$, where P_f^i is the share of patents of firm f in technology class i . This vector also characterizes the

²²Many thanks to Roberto Samaniego for sharing the firm-level R&D intensity data. The same method is also used in Rajan and Zingales (1998) and Ilyina and Samaniego (2011). Outliers (the top and bottom 1% of observations) in the sample are removed to reduce the impact of possible measurement error. The relationship does not change much when we use mean ratio instead of median.

firm’s knowledge distribution. A firm’s overall technology applicability measure, TA_f , is calculated as the weighted geometric mean of the applicability of its technologies: $TA_f = \prod_{i \in J} (a^i)^{P_f^i}$.²³ To measure multi-technology patenting (or technology scope), we count the number of distinct technology classes in which firm has patented.

Observation 1: *Larger firms (measured by sales and number of employees) innovate more and cover more patent classes.*

Table I shows that standard measures of firm size (total sales and number of employees) are highly correlated with firms’ patent stock and the number of technology classes using the Compustat-Patent data matched by Hall, et al. (2005). Especially, a firm with a larger patent stock also conducts R&D in a greater number of patent categories (the correlation equals 0.95). The correlation between patent stock and sales (size of employees) is also as high as 0.69 (0.66).

Table I: Correlation between Patent Stock, Patent Scope and Firm Size

Variables (in log)	No. of Patents	No. of Tech Categories	Sales	No. of Employees
No. of Patents	1			
No. of Tech Categories	0.951	1		
Sales	0.692	0.711	1	
No. of Employees	0.662	0.704	0.952	1

Observation 2: *Firms with more patents (or more patent classes) are more concentrated in highly applicable technologies.*

Observation 3: *Yet, their recent patents are in the less applicable ones.*

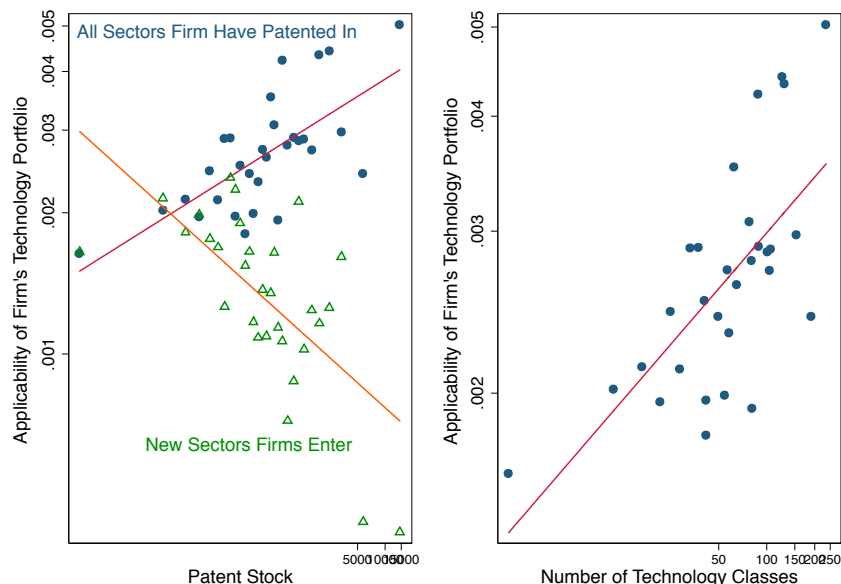
Figure V illustrates the scale dependence in firms’ patent distribution and entry pattern. The left panel plots firms’ technology applicability, TA_f , against their patent stock, distinguishing sectors a firm entered in 2000 (the downward sloping fitted line) from sectors in which the firm has previously patented (the upward sloping line).²⁴ The right panel plots firms’ technology applicability against numbers of technological areas in which the firms are engaged in research (i.e. the firm’s technology scope). Firms’ patent stock and numbers of technology classes are each divided into 30 bins, and each figure presents the variable of interest according to the bin.

Two observations stand out. First, the firm with a higher patent stock (left panel) or broader technological scope (right panel) tends to innovate more in highly applicable technologies. This

²³We use geometric mean instead of arithmetic mean is because (a) technology applicability (a^i) has highly skewed distribution and taking the log of $\{a^i\}$ generates more dispersed distribution; (b) statistically geometric mean is less affected by outliers.

²⁴A sector is new to a firm if the firm has not innovated in that sector before. The full data set expands from 1901 to 2006, thus, provides a good sample for identifying new sectors.

Figure V: Firm's Technology Applicability, Patent Stock and Multi-Technology Patenting



Notes: Y-axis measures the (weighted) average applicability of the firm's patent portfolio, TA_f . Firms are divided into 30 bins according to their patent stocks (left panel) or their numbers of technology classes (right panel). Each observation corresponds to an average firm in the size bin. Both horizontal and vertical axes are in log scale. Data source: NBER Patent Data, 2006 edition.

observation, however, is sharply reversed when focusing on the firm's recent patent classes: the new sectors's applicability is negatively related to firm size (measured either by patent stock or the number of classes). Second, across firms of various sizes, the new sectors entered by a given firm tend to be less applicable relative to the existing sectors, except for the very small firms (i.e. the observations that identify new sectors lie below the observations of all sectors).

In Appendix A.3, we provide further evidence on the relationship between a firm's patent distribution, patent stock and multi-technology patenting using fixed-effect panel regressions. We also show that these results are robust to different levels of disaggregation.²⁵

Observation 4: *Controlling for the initial patent stock, firms whose technologies are more applicable innovate faster.*

We find that the applicability of a firm's initial technology mix matters for its subsequent innovation rate. When regressing firms' subsequent innovation rate on their initial knowledge

²⁵Results are similar at 42 3-digit Standard Industry Classification (SIC) industry level, or at the International Patent Classification (IPC) level which is based on 977 technology classes.

applicability and patent stock, we obtain the following result:²⁶

$$\frac{\Delta S_{f,1}}{S_{f,0}} = \frac{3.46}{(0.229)} - \frac{.50}{(0.009)} \ln(S_{f,0}) + \frac{.27}{(0.078)} \ln(TA_{f,0}), \quad R^2 = 0.16$$

where $\Delta S_{f,1}$ is the number of new patents by firm f in the period 1990-2000, $S_{f,0}$ is firm f 's (accumulated) patent stock in 1990 and $TA_{f,0}$ is the firm's technology applicability in 1990. The positive coefficient on the term $TA_{f,0}$ indicates that after controlling for firms' initial knowledge capital, firms which concentrate more on applicable technologies have higher innovation rates. Although not the focus of our model, the estimation result also shows that firms with larger initial knowledge stock tend to innovate more slowly in the subsequent period (the coefficient on the term $\ln(S_{f,0})$ is negative). This could reflect the decreasing return of learning from others to the size of its existing knowledge capital.²⁷

In Appendix A.3, we present estimation results from panel regressions, controlling for firm sizes (measured by employment), firm fixed effect and selection bias using Heckman two-step procedure. In addition, we also investigate the extensive margin (innovation in new technological areas) and the intensive margin (innovation in existing technological areas) of firm innovation separately. We find that higher initial technology applicability leads to higher innovation rates (overall, intensive and extensive), as well as higher survival probability. This implies that a central location on the technology space enhances firm innovation by providing better prerequisite knowledge for future expansion.

3 Model

Our model extends the previous literature on firm innovation and growth (especially, Klette and Kortum, 2004; henceforth, KK) to a multi-sector environment. It regards product innovation as a process of generating new varieties in different sectors by applying existing knowledge in *all* sectors. Thus, the model is built on the tradition of variety expanding models (e.g., Romer 1990; Grossman and Helpman 1991a; Jones 1995). Recently, Balasubramanian and Sivadasan (2011) provides strong empirical evidence showing that firm patenting is associated with firm growth through the introduction of new products.²⁸ The strong link between a firm's patent stock and its

²⁶We also investigate quality-adjusted innovation rates, which are measured by the growth rates of the forward-citation-weighted number of patents. When adjusted by the number of inward citations, larger firms' growth rates drop even faster, because the number of inward citations per patent decreases with firm size in both the extensive margin (number of classes) and the intensive margin (number of patents within the class).

²⁷Related to this observation, using firm-level data, Akcigit (2009) also finds that firm growth is negatively related to firm size.

²⁸Earlier evidence cited by Scherer (1980) also shows that firms allocate 87% of their research outlays to product improvement and developing new products and the rest to developing new processes.

technology scope also suggests that it may be important to consider firm scope as the source of heterogeneity across innovating firms.²⁹

In linking the model to the data, we interpret our *sector* as corresponding to different *technology classes* in the patent data, while *varieties* within a sector map into *patents* granted in that technology class. We also equate patents with innovations (blueprints).

We first describe the goods demand and firm's static production decision, following the standard setup in the variety-expanding literature, in Section 3.1. Section 3.2 sets out the knowledge production process. We then introduce the main departure of our model from the existing literature: the dynamic multi-sector innovation decisions of firms in Section 3.3 and firms' sectoral selection in Section 3.4. The aggregate equilibrium conditions and equilibrium definitions are given in Section 3.5 and 3.6, respectively.

3.1 Goods Demand and Production

Demand The economy is populated by a unit measure of identical infinitely-lived households. Households do not value leisure, and order their preferences over a lifetime stream of consumption $\{C_t\}$ of the single final good according to

$$U = \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta}}{1-\eta} \quad (1)$$

where β is the discount factor and η is the risk-aversion coefficient. A typical household inelastically supplies a fixed unit of labor, L , which the household can allocate to work as production workers, researchers or workers in the licensing (or lab-rental) industry. Households have access to a one-period risk-free bond with interest rate r_t and in zero aggregate supply. Maximizing their lifetime utility subject to an intertemporal budget constraint requires that consumption evolves according to

$$\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\eta} \frac{P_t}{P_{t+1}} (1 + r_t) = 1, \quad (2)$$

where P_t is the price of the final good.

There are three types of goods in the economy: a final consumption good, sectoral goods and sectoral-differentiated varieties. To concentrate on the heterogeneity in knowledge spillovers across sectors, we abstract from other possible sources of sectoral heterogeneities, such as expenditure shares, elasticities of substitution between varieties and within-sector cross-firm knowledge spillover intensities. The final good is produced by combining quantities of K different sectoral intermediate

²⁹Also in KK and Bernard Redding and Schott (2011) firms are heterogeneous in terms of their product scopes.

goods $\{Q_t^i\}$ according to a Cobb-Douglas production function

$$\log Y_t = \sum_{i=1}^K s^i \log(Q_t^i), \quad (3)$$

where, $s^i = 1/K$ captures the share of each sector in production of the final good. Without physical capital in the closed-economy model, the final good is only used for consumption: $C_t = Y_t$.

At any moment, each sector contains a set of varieties that were invented before time t . In particular, we represent the set of varieties in sector i available on the market by the interval $[0, n_t^i]$. Sector i good is aggregated over these n_t^i number (measure) of differentiated goods that are produced by individual monopolistically competitive firms

$$Q_t^i = \left[\int_0^{n_t^i} (x_{k,t}^i)^{\frac{\sigma-1}{\sigma}} dk \right]^{\frac{\sigma}{\sigma-1}}, \quad i = 1, 2, \dots, K, \quad (4)$$

where $x_{k,t}^i$ is the consumption of variety k in sector i and $\sigma > 1$ is the elasticity of substitution between differentiated goods of the same sector i . Each new variety substitutes imperfectly for existing ones, and the firm which develops it exploits limited monopoly power in the product market.

The associated final good price is $P_t = B \prod_i^K (P_t^i)^{s^i}$, where B is some constant consistent with the Cobb-Douglas specification in (3) and sectoral price index, P_t^i is given by

$$P_t^i = \left[\int_0^{n_t^i} p_{k,t}^{1-\sigma} dk \right]^{\frac{1}{1-\sigma}}. \quad (5)$$

These aggregates can then be used to derive the optimal consumption for sector- i goods and for individual variety k in sector i using

$$Q_t^i = \frac{s^i P_t Y_t}{P_t^i}, \quad (6)$$

$$x_{k,t}^i = \left(\frac{p_{k,t}^i}{P_t^i} \right)^{-\sigma} Q_t^i. \quad (7)$$

Production Firms undertake two distinct activities: they create blueprints for new varieties of differentiated products, and they manufacture the products that have been invented. The firm inventing a new variety is the sole supplier of that variety. As the focus is upon firms' innovation activities, the production side of the model is kept as simple as possible. We assume that each differentiated good is manufactured according to a common technology: to produce one unit of any

variety requires one unit of labor, $y_{ft}^i = l_{ft}^i$, $\forall i, f$.

Without heterogeneity in supply and demand, all varieties in the same sector are completely symmetric: they charge the same price and are sold in the same quantity. The firm producing variety k in sector i faces a residual demand curve with constant elasticity σ specified in (7).³⁰ Wage is normalized to one. This yields a constant pricing rule $p_{k,t}^i = \frac{\sigma}{\sigma-1}$, $\forall k, i$ and t . Thus the sectoral price, $P_t^i = \frac{\sigma}{\sigma-1}(n_t^i)^{\frac{1}{1-\sigma}}$, decreases with the total number of varieties in that sector as $\sigma > 1$.

Combining the pricing rules with (5) and (7), we derive the *total* profit in the *product market* in sector i (aggregated over all varieties produced by different firms) as a constant share of GDP, $P_t Y_t$:

$$\pi_t^i = \int_0^{n_t^i} \frac{p_{k,t}^i x_{k,t}^i}{\sigma} dk = \frac{s^i P_t Y_t}{\sigma}. \quad (8)$$

The total demand for production labor in sector i is

$$L_{p,t}^i = \int_0^{n_t^i} x_{k,t}^i dk = \frac{\sigma-1}{\sigma} s^i P_t Y_t. \quad (9)$$

3.2 Knowledge Creation

There is a continuum of firms, each developing new varieties and producing in multiple sectors. A firm at time t is defined by a vector of its differentiated products in all sectors,

$$\mathbf{z}_{f,t} = (z_{f,t}^1, z_{f,t}^2, \dots, z_{f,t}^K)',$$

where $z_{f,t}^i \geq 0$ is the number of differentiated sector- i goods produced by firm f at time t . To add new varieties to its set, a firm devotes a given amount of labor to R&D. Since only the firm inventing the variety has the right to manufacture it, $\mathbf{z}_{f,t}$ also characterizes the distribution of the firm's private knowledge capital across sectors.

Let \mathcal{J} be the set of all sectors. Then $|\mathcal{J}| = K$, and $\mathcal{S}_{f,t} \subseteq \mathcal{J}$ denote the subset of sectors in which firm f produces at time t , i.e. $\mathcal{S}_{f,t} = \{i: \text{s.t. } z_{f,t}^i > 0\}$. Let $\mathcal{F}_{i,t} = \{f: \text{s.t. } z_{f,t}^i > 0\}$ denote the set of firms that produce in sector i . Then $n_t^i = \int_{f \in \mathcal{F}_{i,t}} z_{f,t}^i df$.

Consider a firm f in sector i with a stock $z_{f,t}^i$ of private knowledge at time t . For simplicity, we assume knowledge never depreciates. The sectoral knowledge of firm f , thus, accumulates over time according to

$$z_{f,t+1}^i = z_{f,t}^i + \Delta z_{f,t}^i, \quad (10)$$

³⁰To make the analysis more tractable, we follow Hopenhayn (1992) and Klette and Kortum (2004) by assuming that each firm is relatively small compared to the entire sector.

New sectoral knowledge (or new varieties), $\Delta z_{f,t}^i$, is generated based on an innovation production function, using the firm's R&D input and accessible knowledge stock in *all* sectors and is subject to idiosyncratic innovation shocks. Since knowledge spillovers across sectors are heterogenous, we decompose firm's sectoral R&D investment according to its *knowledge source sector*.³¹ For clarity, we introduce the following notation: $R_f^{i \leftarrow j}$ denotes a firm's R&D input when utilizing sector j 's knowledge to generate new knowledge (invent new blueprints) in sector i . The arrow indicates the direction of knowledge flow (when necessary): i represents the sector that the firm is applying the knowledge to—the application (or target) sector and j is the sector that the firm is adopting knowledge from—the source sector.³² Thus, each innovation activity is defined by two sectors.

The new sector- i knowledge created by firm f summarizes innovation output in different R&D activities, each utilizing a different type of source knowledge j , $j \in \mathcal{J}$. The firm can use its private source knowledge capital to innovate, or the public knowledge to imitate. One of the central notions of our paper is that the productivity of innovation inputs depends on the elements of the knowledge diffusion matrix, $A = (A^{i \leftarrow j})_{(i,j) \in \mathcal{J} \times \mathcal{J}}$, which is taken as exogenous by firms.³³ Specifically, new knowledge in i is produced based on a Cobb-Douglas combination of innovation productivity $((A^{i \leftarrow j})_{j \in \mathcal{J}})$, the firm's current R&D investment $((R_{1f}^{i \leftarrow j})_{j \in \mathcal{J}}$ in innovation and $(R_{2f}^{i \leftarrow j})_{j \in \mathcal{J}}$ in imitation) and its stock of source knowledge (private knowledge capital, $(z_f^j)_{j \in \mathcal{J}}$ and public knowledge capital $(\bar{z}^j)_{j \in \mathcal{J}}$):³⁴

$$\Delta z_{f,t}^i = \sum_{j=1}^K \left[A^{i \leftarrow j} \left(\bar{z}_t^i R_{1f,t}^{i \leftarrow j} \right)^\alpha \left(z_{f,t}^j \right)^{1-\alpha} \varepsilon_{1f,t}^{ij} + A^{i \leftarrow j} \left(\bar{z}_t^i R_{2f,t}^{i \leftarrow j} \right)^\alpha \left(\theta \bar{z}_t^j \right)^{1-\alpha} \varepsilon_{2f,t}^{ij} \right] \quad (11)$$

where α is the share of R&D in the innovation production. We explain the elements of this production function in turn as follows.

First, similarly to KK, we assume that the production function of each innovation activity is constant returns to scale. In addition, the researchers' efficiency is assumed to be proportional to the average knowledge per firm in the innovating sector, \bar{z}_t^i , thus the effective R&D is given by $\bar{z}_t^i R_{kf,t}^i$, $k = 1, 2$. This assumption keeps the total number of R&D workers constant in the

³¹Firms have to devote a certain amount of time digesting and adopting knowledge in one sector to apply it to another.

³²When $i = j$, it captures the within-sector knowledge spillovers.

³³It might be true that technologies advance over time and the interaction between one another evolves, forming a dynamic network instead of a static one. Also, these relationships of complementarity may be hard to predict and not necessarily visible or well understood by innovators. Here, we intentionally choose to concentrate on the implications of very 'deep', time-invariant characteristics of technological linkages on firm's innovation and leave the study of dynamic knowledge network formation to future work, as we clearly view it as a necessary first step.

³⁴We use additive instead of multiplicative function to combine the knowledge capital in different sectors, because the additive function of firm size in different sectors can generate Pareto distribution of $\{z_{f,t}^i\}$ in each sector i ; any linear combination of $\{z_{f,t}^i\}$ also follows Pareto distribution according to Kesten(1973). Besides, firm value function is linear in $\{z_{f,t}^i\}$ under additive knowledge production function, which makes the model more tractable.

stationary equilibrium while the number of goods grows. Also, as will be explained later in Section 4.3, it helps to remove the ‘scale effect’ from the model—that is, the endogenous growth rate of the economy is independent of the population size.

Second, in the process of developing new blueprints in sector i , a firm utilizes all existing knowledge at its disposal: its private knowledge from every sector $j \in \mathcal{S}_{f,t}$, and public knowledge from all sectors. Here, we assume the size of the public knowledge pool is proportional to the average knowledge per firm in sector j , \bar{z}_t^j , for the following reasons: When learning from others is costly, each firm is too small to access all stock of knowledge in the whole sector. When firms randomly meet and learn from a limited number of peers, the average knowledge capital *per firm* is a better proxy for the size of public knowledge than the total knowledge stock in that sector.³⁵ θ governs the accessibility of the public knowledge relative to the in-house knowledge.

Third, innovation by its nature includes the discovery of the unknown; therefore, the success of a research project can be uncertain. We assume that firm innovation and imitation are subject to shocks $\varepsilon_{1f,t}^{ij}$ and $\varepsilon_{2f,t}^{ij}$, respectively, which follow the same identical and independent distribution $G(\varepsilon)$ across firm, sector-pairs and time.³⁶ Firms know the distribution of shocks but not their actual realizations before deciding on the optimal R&D input. A series of large negative shocks lead to exit and a series of positive ones cause further expansion. Later we will show that these i.i.d. shocks endogenously generate a Pareto firm size distribution in every sector and in the aggregate economy.

3.3 Firm R&D Decisions

We now determine firms’ R&D effort. A firm may enter sectors freely, but must pay a fixed research cost of $F_{f,t}^i$ (measured in units of labor) every period in order to develop new varieties in a given sector i . This fixed cost, $F_{f,t}^i = F\zeta_{f,t}^i$, has two components: a constant term F that is identical for all firms and all sectors; and a firm-specific idiosyncratic component, $\zeta_{f,t}^i$, which is assumed to be i.i.d. across sectors, firms and time, and satisfies $E\zeta_{f,t}^i = 1$. If a firm does not pay this cost, then it ceases to develop new products in that sector. This continuation cost can be interpreted as a license fee or the financial cost of maintaining a research lab.

The timing works as follows. In each period, a firm first makes a draw of the idiosyncratic cost $\zeta_{f,t}^i$ from an underlying distribution $H(\zeta)$, and then chooses to stay in (or enter) sector i or

³⁵ As shown later, this assumption also helps to ensure that the sectoral growth rate is independent of the number of firms and the total population in the general equilibrium.

³⁶ The mean of $\varepsilon_{kf,t}^{ij}$, $k = 1, 2$ is set to be 1 and bounded by zero such that the innovation rate is always positive. A firm’s market share in sector i increases only if its growth rate beats the average growth rate in the sector. If a firm stops conducting R&D, its market share will shrink to zero eventually. In this way ‘creative destruction’ is embodied in this model.

discontinue this research line. If its expected additional payoff from continuing innovating in that sector is greater than the fixed cost, the firm decides on the optimal R&D investment, financed by issuing equity. After that, firm-specific innovation shocks realize and the firm creates $\Delta z_{f,t}^i$ new blueprints. If the continuation value is lower than the fixed cost, the firm discontinues its research in that sector, sells its blueprints and exits that sector.

Given the assumption of a continuum of firms, in equilibrium there always exists a mass of very large firms that are operating in all sectors, and would never exit any sector.³⁷ We first specify the R&D decision making process of such a large all-sector firm. Since this kind of firm never dies, the per-period fixed cost would not affect the firm's R&D decisions, but simply reduce the firm's value by the present value of future fixed costs: $F_{f,t} + E_t \frac{F_{f,t+1}}{1+r} + E_t \frac{F_{f,t+1}}{(1+r)^2} + \dots = F_{f,t} + \frac{F}{r}$. We can then solve for the all-sector firm's R&D decision problem as if the firm had paid the initial sunk entry cost, and was only concerned about the optimal R&D investment across all sectors.

Since each variety is sold and priced at the same level, the firm f 's market share in sector j can be captured by $\frac{z_{f,t}^j}{n_t^j}$. An all-sector firm that receives a flow of profit $\sum_{j=1}^K \pi_t^j \frac{z_{f,t}^j}{n_t^j}$ in the product market chooses an R&D investment portfolio to maximize its (post-sunk-cost) expected present value $V(\mathbf{z}_{f,t})$, given the interest rate r_t . By spending on R&D, the firm incurs a cost of hiring researchers, whose wage rate is normalized to one. The new blueprints will be turned into products and sold in the next period. The firm's Bellman equation is

$$\max_{(R_{1f,t}^{i \leftarrow j})_{i,j \in \mathcal{J} \times \mathcal{J}}, (R_{2f,t}^{i \leftarrow j})_{i,j \in \mathcal{J} \times \mathcal{J}}} V(\mathbf{z}_{f,t}) = \sum_{j=1}^K \pi_t^j \frac{z_{f,t}^j}{n_t^j} - \sum_{i=1}^K \sum_{j=1}^K \left(R_{1f,t}^{i \leftarrow j} + R_{2f,t}^{i \leftarrow j} \right) + \frac{1}{1+r_t} E[V(\mathbf{z}_{f,t+1})] \quad (12)$$

subject to the knowledge accumulation equation (10) and the incremental innovation production function (11).

This paper only considers the stationary balanced growth path (BGP) equilibrium in which the growth rates of aggregate variables remain constant over time (it is formally defined in Section 3.6). The full characterization of the dynamics of firm value is presented in Appendix (B.1). In the BGP equilibrium, the aggregate profit in the product market at the sector level is constant, *i.e.* $\pi_t^j = \pi^j$ (because the supply of the only production factor L is fixed). The interest rate also remains constant $r_t = r$ and is pinned down by (2). Define the BGP growth rate of the number of varieties in sector i as $\gamma_t^i \equiv n_{t+1}^i / n_t^i$. In Appendix (B.1), we prove that on the BGP, different sectors grow at the same rate, that is $\gamma_t^i = \gamma, \forall i$. The basic intuition is that cross-sector knowledge spillovers keep all sectors on the same track. Therefore, the distribution of the number of varieties (knowledge stock) across sectors is stable and invariant: $n_t^i / n_t^j = n^i / n^j$. Also, the number (mass)

³⁷An alternative interpretation is that there exists a large research institute which never dies and is willing to purchase new blueprints at their market value.

of firms (M_t^i) in every sector in the stationary BGP does not change over time, i.e. $M_t^i = M^i$ and $\bar{z}_t^i/n_t^i = 1/M^i$, $\forall i$. Notice that in such a BGP equilibrium, economy-wide or sector-wide aggregates grow at constant rates, but firm growth rates, entry and exit into different sectors are heterogenous.

The linear form of the Bellman equation (12) and the constant returns to scale (Cobb-Douglas) innovation technology allow us to derive closed form solutions for the above optimization problem. Define $\rho \equiv \frac{1}{1+r} \frac{1}{\gamma}$. It is easy to verify that in the stationary BGP equilibrium, the firm's value is a *linear* aggregate of the value of its knowledge capital in all sectors,

$$V(\mathbf{z}_{f,t}) = \sum_{i=1}^K \left(v^i \frac{z_{f,t}^i}{n_t^i} + u^i \right),$$

where v^i is the market value of total knowledge capital in sector i , which is time-invariant on BGP and is given by

$$v^i = \frac{1}{(1-\rho)} (\pi^i + \sum_{j=1}^K \omega^{j \leftarrow i}), \quad (13)$$

and $\omega^{j \leftarrow i}$ captures the *application value* of sector i 's knowledge stock to innovation in sector j ,

$$\omega^{j \leftarrow i} = \frac{1-\alpha}{\alpha} \frac{n^i}{n^j} (A^{j \leftarrow i} \alpha \rho v^j)^{\frac{1}{1-\alpha}} (M^j)^{\frac{\alpha}{\alpha-1}}. \quad (14)$$

We refer to u^i as the rent from public knowledge (imitation), measured by the aggregate application value generated by all sectors to sector i .

$$u^i = \left(1 + \frac{1}{r} \right) \sum_{j=1}^K \omega^{i \leftarrow j} \left(\frac{\theta \bar{z}_t^j}{n_t^j} \right) = \left(1 + \frac{1}{r} \right) \sum_{j=1}^K \frac{\theta \omega^{i \leftarrow j}}{M^j} \quad (15)$$

Clearly from (13) and (14), solving for the equilibrium price of sectoral knowledge capital is an *iterative process*: the knowledge value of any given sector depends upon the knowledge value of all other sectors. Overall, the relative prices of knowledge capital in different sectors (v^i/v^j) are determined by the exogenous fundamental linkages between sectors (captured by $A^{i \leftarrow j}$) and other general equilibrium conditions.

The interpretations for (13) (15) and (14) are intuitive. (13) shows that the value of all the blueprints in sector i , v^i , is not limited to the *direct* profit return ($\frac{\pi^i}{1-\rho}$)—but also depends upon its *indirect* capital value captured by its contribution to future innovations in all K sectors ($\frac{\sum_{j=1}^K \omega^{j \leftarrow i}}{1-\rho}$). (14) implies that the knowledge application value of j to i is larger when sector j 's knowledge stock is relatively more abundant (higher n^j/n^i), or when the knowledge in target sector i is more valuable (higher v^i), or when the knowledge spillovers from j to i is stronger (larger $A^{i \leftarrow j}$), or when

sector i is less competitive (lower M^i). (15) implies that when public knowledge is easier to access (higher θ) or when knowledge in other sectors is more applicable and more valuable (higher $\omega^{i \leftarrow j}$), the rent from external knowledge is higher.

The *optimal R&D investment* associated with applying sector- j knowledge to sector i is

$$R_{f,t}^{i \leftarrow j} \equiv R_{1f,t}^{i \leftarrow j} + R_{2f,t}^{i \leftarrow j} = \frac{\alpha}{1 - \alpha} \omega^{i \leftarrow j} \frac{z_{f,t}^j + \theta \bar{z}_t^j}{n_t^j}. \quad (16)$$

A firm scales up its R&D investment in proportion to the application value of sector j 's knowledge to sector j , $\omega^{i \leftarrow j}$, and its (normalized) accessible knowledge capital.

We now turn to address the innovation decisions of firms that have only entered a subset of sectors. We assume that the knowledge capital market is efficient.³⁸ Under this assumption, the all-sector firms would bid up the price of each blueprint in every sector, because they are the most diversified firms and can fully internalize and utilize the new knowledge in every sector. As a result, the market price of a blueprint is equivalent to the price that an all-sector firm is willing to pay, which is given by $\frac{v^i}{n_t^i}$ at time t . Importantly, we assume that upon exit from a specific sector, a firm can sell all its blueprints at the market price and thus does not lose the value of its private knowledge. As long as there exist such potential buyers at any given time, the market price of knowledge capital will be bid up to its marginal value for an all-sector firm. Therefore, a small firm, upon entering a sector, takes the price of blueprints in different sectors as given and makes decisions on its optimal R&D investment portfolio. The solution would be the same as in (16).

3.4 Sectoral Entry and Exit

As explained earlier, to continue its research in sector i , firms incur a period-by-period fixed continuation cost. If a firm does not pay this cost, then it ceases to develop new products and has to sell its blueprints and exit the sector. Under free entry, a firm drawing a cost of $F^i \zeta_{f,t}^i$ will continue its research in sector i or enter this sector if the additional value created by this action can cover the cost. That is

$$F^i \zeta_{f,t}^i \leq - \sum_{j=1}^K R_{f,t}^{i \leftarrow j} + \frac{1}{1 + r} E_t[V(\dots, z_{f,t}^i + \Delta z_{f,t}^i, \dots) - V(\dots, z_{f,t}^i, \dots)]. \quad (17)$$

³⁸The efficient knowledge capital market assumption significantly simplifies the analysis. Otherwise, firms with small knowledge scope would not be as motivated to conduct R&D, since they could not internalize intersectoral knowledge spillovers as complete as an all-sector firm. Without an efficient knowledge capital market, the price of each blueprint will be inventor-specific and tracking the values of all blueprints of all firms is almost computationally impossible.

The effort creates additional value of $v^i \Delta z_{f,t}^j / n_{t+1}^i$ for the firm in the next period, where v^i is given in (13). Combining (13) and (16) we can rewrite (17) as

$$F^i \zeta_{f,t}^i \leq - \sum_{j=1}^K R_{f,t}^{i \leftarrow j} + \frac{1}{1+r} \left(\frac{v^i E_t \Delta z_{f,t}^i}{n_{t+1}^i} \right) = \sum_{j=1}^K \omega^{i \leftarrow j} \frac{(z_{f,t}^j + \theta \bar{z}_t^j)}{n_t^j} = \sum_{j=1}^K \omega^{i \leftarrow j} \frac{z_{f,t}^j}{n_t^j} + \frac{r}{1+r} u^i. \quad (18)$$

The last equation says that a potential entrant to sector i (i.e. $z_{f,t}^i = 0$) can apply its private and public knowledge capital from *all* the related sectors to invent new products in the entering sector.³⁹ Therefore, in this multi-sector model, firms with different knowledge mix $(z_{f,t}^j / n_t^j)_{j \in \mathcal{S}_{f,t}}$ self-select into different sectors. Given the definition of $\omega^{i \leftarrow j}$ in (14), large positive elements in the i^{th} row of the knowledge diffusion matrix and higher value of sector i 's knowledge, v^i , attract more potential entrants. On the other hand, a larger number of existing products, n^i , or more incumbent firms, M^i , deter entry.

New Firms There is a large pool of prospective new firms in the economy. A new firm—a firm which has invented no blueprints in *any* sector ($z_{f,t}^i = 0 \ \forall i$)—enters the economy by starting from the sector where the fixed cost can be covered by imitation (the application value of the existing set of public knowledge capital). The free entry condition for the newborn firm implies that it will first enter the sector i that offers the largest benefit of knowledge spillovers minus the fixed cost:

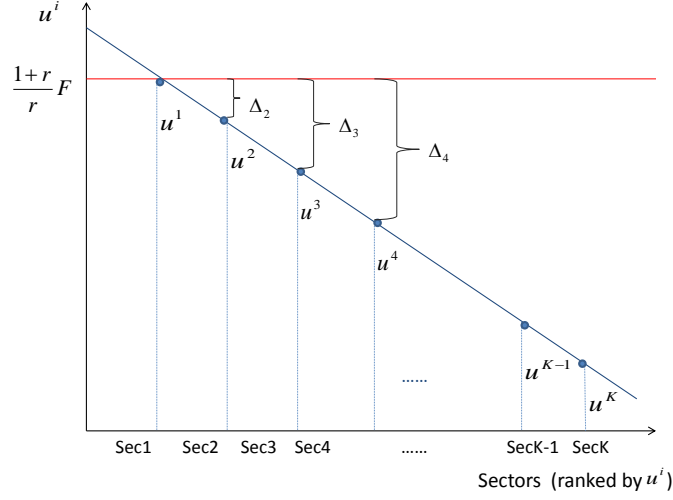
$$i = \arg \max_i \left\{ \frac{r}{1+r} u^i - F \zeta_f^i \right\}. \quad (19)$$

Since firms have different draws of sector-specific fixed cost ζ_f^i , the first sectors that new firms enter may not be the same.

Sequential Sectoral Entry The sectoral entry condition (18) along with equation (19), imply that firms enter different sectors *sequentially*: they start developing new varieties in a sector that offers the largest public knowledge externality, building up private knowledge and then venturing into other sectors using its accumulated knowledge. The sequential sectoral entry can be better explained using Figure VI. Suppose sectors are ranked by their externality of public knowledge, and $u^1 > u^2 > \dots > u^K$. If firms all draw the same fixed cost F , every new firm enters sector 1 first. Entry stops when the net value of entry is zero. Next, in order to enter more sectors, the firm needs to accumulate more private knowledge to fill up the gap between the entry cost and the free knowledge externality provided by the public knowledge, that is $\Delta_2, \Delta_3, \dots$, etc. Since firms

³⁹Note that significantly different from previous models of entry, prior to entry, potential entrants are not identical; they *differ* in terms of their knowledge mix $(z_{f,t}^i)_{i \in \mathcal{S}_{f,t}}$.

Figure VI: Determination of Firm's Entry into Multiple Sectors



are facing idiosyncratic shocks to innovation and fixed costs, not all firms follow the exact same path expanding across the technology space. Yet, their entries are all *path-dependent*: depending on where they have entered in the past, the intersectoral knowledge linkages determines the next optimal step.

Exit A firm stops inventing new varieties in sector i if the fixed cost is higher than the expected benefit of continuing R&D. A firm that discontinues its R&D in a sector can sell its blueprints (knowledge capital) in this sector to an all-sector firm for the price of v^i/n^i per variety. Once the patent is sold it can no longer be used it to invent in other sectors.⁴⁰

3.5 Aggregate Conditions

The population supplies L units of labor services at every period and they are allocated in three areas: production workers allocated in different sectors, researchers and workers who are engaged in applying entry licenses. Formally, the labor market clearing condition is:

$$L = \sum_{i=1}^K L_{p,t}^i + \sum_{i=1}^K \sum_{j=1}^K \int_{f \in \mathcal{F}_i \cap \mathcal{F}_j} R_{f,t}^{i \leftarrow j} df + \sum_{i=1}^K \int_{f \in \mathcal{F}_i} F^i \zeta_{f,t}^i df \quad (20)$$

⁴⁰ Alternatively, it can potentially still produce and sell their previously invented varieties in the product market, as well as apply its accumulated knowledge capital in the exiting sector to invent in other related areas. In equilibrium, these two options generate exactly the same value; thus, the firm is indifferent in keeping the blueprints or not. The reason is because the discounted value of future payoffs associated with the body of knowledge is already fully priced in the value of the sectoral knowledge, v^i . A firm completely exits sector i if it is hit by a series of negative shocks such that $z_f^i \leq 0$ according to its knowledge accumulation in (11).

Using (9) and (16) we can rewrite (20) in the stationary BGP equilibrium as:

$$L = \sum_{i=1}^K \left[\frac{\sigma-1}{\sigma} sL + \alpha\rho(\gamma-1)v^i + F^i M^i \right]. \quad (21)$$

In this economy, the household owns all the firms and finances all the potential entrants. Given an interest rate r , every period the household gets net income $r \sum_i [v^i + (u^i - \frac{1+r}{r}F)M^i]$ from investing in firms.⁴¹ The household's total income is

$$PY = L + r \sum_{i=1}^K \left[v^i + (u^i - \frac{1+r}{r}F)M^i \right] \quad (22)$$

Therefore, according to (8) the sectoral profit π^i in the stationary BGP equilibrium is indeed a constant. Following (2), the stationary BGP interest rate is determined by

$$1 = \beta(1+r)\gamma^{\frac{\eta-1}{1-\sigma}} \quad (23)$$

3.6 Equilibrium Definitions

Definition 1 *An equilibrium is defined as time paths of aggregate consumption, output and price $\{C_t, Y_t, P_t\}_{t=0}^\infty$ that satisfy (21),(22),(23) and the goods market clearing condition $C_t = Y_t$; time paths of consumption levels, numbers of varieties, measure of firms, the total value of blueprints in different sectors $\{n_t^i, M_t^i, Q_t^i, v^i\}_{i=1,\dots,K,t=0}^\infty$ that satisfy (5) (6) (13) (18) and (21); time paths of R&D investment, sectoral innovation (production) and prices by different firms $\{R_{f,t}^{i \leftarrow j}\}_{i,j=1,\dots,K,f \in \mathcal{F}_{j,t},t=0}^\infty$ $\{z_{f,t}^i, p_{f,t}^i\}_{i=1,\dots,K,f \in \mathcal{F}_{i,t},t=0}^\infty$ that maximize discounted present firm value, that is, satisfy (10) (11) and (16) ; time paths of firm's sectoral entry and exit decisions that satisfy (18) and time paths of wage and interest rates $\{w_t, r_t\}_{t=0}^\infty$ that satisfies (2) and $w_t = 1$.*

Definition 2 *A balanced growth path (BGP) is an equilibrium path in which output, consumption and innovation grow at constant rates.*

Definition 3 *A stationary BGP equilibrium is a BGP in which the distribution of normalized firm sizes is stationary in every sector.*

Throughout the paper, we analyze a stationary BGP equilibrium defined in the section above. In Section 4.1 we show that our model endogenously generates stationary firm size distribution that converges to a Pareto distribution when the number of firms is extremely large.

⁴¹Equivalent to getting dividend as profit and capital gains.

4 Stationary BGP Equilibrium

4.1 Firm Size Distribution

In a typical firm's life span, the firm starts from a relatively highly applicable center sector. After accumulating enough background knowledge, a small firm with a sequence of good draws of innovation shocks can expand into related sectors along the inter-sector knowledge linkage network. After several rounds of entry selection, only a few large, multi-sector firms can reach the edge of the technology space.

Since varieties in the same sector are produced at the same quantity, the normalized firm size in sector i for firm f can be given by $\tilde{z}_{f,t}^i = z_{f,t}^i / n_t^i$. Putting (10), (11) and (16) together yields the following firm size dynamics:

$$\tilde{\mathbf{z}}_{f,t+1} = \Phi_{\mathbf{f},\mathbf{t}} \tilde{\mathbf{z}}_{f,t} + \Psi_{\mathbf{f},\mathbf{t}} \mathbf{b}, \quad (24)$$

where the K -dimensional vector $\tilde{\mathbf{z}}_{f,t} \equiv (\tilde{z}_{f,t}^1, \dots, \tilde{z}_{f,t}^K)$, the constant vector $\mathbf{b} \equiv (\theta/M^1, \dots, \theta/M^K)$ and the $\{i, j\}^{th}$ elements of the $K \times K$ matrices $\Phi_{\mathbf{f},\mathbf{t}}$ and $\Psi_{\mathbf{f},\mathbf{t}}$ are given by $\phi_{f,t}^{ij}$ and $\psi_{f,t}^{ij}$ respectively:

$$\phi_{f,t}^{ij} = \frac{1}{\gamma} \left(1_{\{i=j\}} + \xi^{ij} \varepsilon_{1f,t}^{ij} \right), \quad \psi_{f,t}^{ij} = \frac{\xi^{ij} \varepsilon_{2f,t}^{ij}}{\gamma}.$$

where $\xi^{ij} = \frac{\omega^{ij}}{(1-\alpha)\rho v^i}$, $1_{\{i=j\}}$ is one if $i = j$ and zero otherwise.

According to Kesten (1973), (24) implies that firm size distribution (in each sector and in the whole economy) converges in probability to a Pareto distribution in the upper tail.⁴² The existence of public knowledge plays an important role in attenuating the size dispersion generated by idiosyncratic innovation shocks.

In Section 5, we simulate the model economy with a large number of firms and show that the firm size distribution generated by the model indeed follows Pareto tail distribution.

4.2 Heterogenous R&D Intensities Across Sectors

In this section, we study the sectoral R&D intensity (R&D expenditure as a fraction of sales),

$RI^i \equiv \frac{1}{s^i PY} \sum_{j=1}^K \int_{f \in \mathcal{F}_i \cap \mathcal{F}_j} R_f^{i \leftarrow j} df$. Based on (16), our model predicts that sectoral R&D resources

⁴²The firm size distribution in sector i can be characterized by the distribution of $\mathbf{x}\tilde{\mathbf{z}}_f$, when $\mathbf{x} = (0, 0, \dots, 1, \dots, 0)$ with the i^{th} element being one. Similarly, when $\mathbf{x} = (\frac{1}{K}, \frac{1}{K}, \dots, \frac{1}{K})$, the distribution of $\mathbf{x}\tilde{\mathbf{z}}_f$ captures the firm size distribution in the whole economy. Since power law is conserved under addition and multiplication, the overall firm size distribution in the aggregate economy is also Pareto. For more detailed discussion and application of Kesten (1973), see Gabaix (2009). Luttmer (2007) provides a state-of-art model for firm size distribution, where firms receive an idiosyncratic productivity shock at each period and firm exit provides a natural lower bound for the distribution. Cai (2012) studies how innovation and imitation affects firm size distribution using a one-sector model and provides more explanations in this context.

are allocated according to the sectoral knowledge value (formally derived in Appendix B.2):

$$\frac{RI^i}{RI^j} = \frac{v^i}{v^j} \quad (25)$$

Therefore, any policies that distort the relative knowledge value, v^i/v^j , also cause misallocation of research investment across sectors. Recall that $v^i = (1 - \rho)^{-1}(\pi + \sum_{j=1}^K \omega^{ij})$. (25) implies that R&D intensity in sector i increases with $\sum_{j=1}^K \omega^{ij}$ —which captures the ‘technology opportunities’, one of the main factors identified in the empirical studies as being the potential determinant of different research intensity across sectors (see Ngai and Samaniego, 2011). This equation is also in line with the observation documented in Section 2.1: R&D intensity and knowledge applicabilities across sectors are positively correlated in the data.

4.3 Aggregate Innovation Rate

The number of varieties in sector i evolves according to $n_{t+1}^i = (n_t^i + \int_{f \in \mathcal{F}_i} \Delta z_{f,t}^i df)$. Define $\tau^{i \leftarrow j}$ as the fraction of sector j ’s knowledge that is actually *utilized* in innovation in sector i , i.e. $\tau^{i \leftarrow j} = \frac{\int_{f \in \mathcal{F}_i} (z_f^j + \theta \bar{z}^j) df}{n^j} \leq 1 + \theta$. On the BGP, all sectors innovate at the same rate. Based on (11) we derive the (gross) growth rate of the number of varieties in the whole economy as

$$\gamma = 1 + \frac{1}{(1 - \alpha)\rho} \sum_{j=1}^K \frac{\omega^{i \leftarrow j} \tau^{i \leftarrow j}}{v^i}. \quad (26)$$

Consider a special case of log utility function ($\eta = 1$). Combining (26) with (13) and (14), we obtain

$$\gamma = (1 - \beta) \left[(1 - \alpha)\beta \frac{\sum_i \sum_j \omega^{i \leftarrow j} + \sum_i \pi^i}{\sum_i \sum_j \omega^{i \leftarrow j} \tau^{i \leftarrow j}} - 1 \right]^{-1}. \quad (27)$$

This equation has three implications. First, everything else being the same, an increase in knowledge linkages across sectors enhances growth (because $\omega^{i \leftarrow j}$ increases). Second, in the presence of fixed costs, not every firm operates in every sector. Therefore, the fraction of knowledge that is utilized across sectors, $\tau^{i \leftarrow j}$, is strictly less than $1 + \theta$ (when knowledge spillovers are complete) $\forall i, j$. Hence, this equation implies that sectoral entry costs reduce the aggregate innovation rate in the economy by blocking the knowledge circulation across sectors. Third and more subtly, when the variance of firm-specific idiosyncratic fixed costs, σ_ζ , rises, firms’ sectoral selection decisions become more random and less driven by the ‘fundamental’ factors—knowledge linkages between firms’ current and entering sectors. Firms with sufficient background knowledge may not be able to conduct research in many sectors, and firms with insufficient background knowledge

may enter many sectors but cannot innovate much. This ‘inefficient’ *sorting* of firms in different sectors—manifest in lower $\sum_i \sum_j \omega^{i \leftarrow j} \tau^{i \leftarrow j}$ —reduces aggregate innovation rate.

4.4 Aggregate Economic Growth

Given that labor supply is fixed, the growth in nominal GDP is zero. However, real growth rate is positive due to the ‘variety effects’. Because of the variety effect and $\sigma > 1$, expansion in varieties is associated with decrease in sectoral prices: $\frac{P_{t+1}^i}{P_t^i} = \left(\frac{n_{t+1}^i}{n_t^i} \right)^{\frac{1}{1-\sigma}}$. Then according to equation (6), $\frac{Q_{t+1}^i}{Q_t^i} = \gamma^{\frac{1}{\sigma-1}}$. The aggregate real output grows at

$$g \equiv \frac{Y_{t+1}}{Y_t} = \prod_{i=1}^K \left(\frac{Q_{t+1}^i}{Q_t^i} \right)^{s^i} = \gamma^{\frac{1}{\sigma-1}}. \quad (28)$$

It is worth pointing out that by assuming the efficiency of R&D workers to be proportional to the average knowledge stock in that sector, we eliminate the ‘scale effects’ of population on economic growth. This can be seen from (26). Both $\omega^{i \leftarrow j}$ and v^i are proportional to the total population in the economy; therefore, the growth rate of varieties is independent of the level of population.⁴³

5 Simulations

We have commented along the way when our model can potentially fit one of the observations in Section 2. Here we simulate the model economy with a large panel of firms (above 30,000) innovating in various (multiple) sectors, to assess the performance of the model in matching the empirical observations that motivated our work.

Allowing for a large number of sectors would give us a more accurate definition of technology classes. It is, however, computationally difficult to calibrate and simulate the economy with a large K .⁴⁴ Therefore, we calibrate the model based on sectors at a less disaggregated level (than in Section 2)—at SIC 2-3 digit industrial classifications provided by the USPTO, which constructs 42 sectors. We calculate the measure of applicability (authority weight) for these 42 sectors using the same method as in Section 2.1. The relevant firm patenting and patent citation data over the 30-year period (1976-2006) are employed to discipline the parameters.

⁴³Jones (1999) first pointed out that the ‘scale effects’ that plague many endogenous growth models are not consistent with empirical evidence.

⁴⁴For example, our empirical evidence shown in Section 2 is based on 428 sectors. It is almost impossible to calibrate all the sectoral parameters for 428 sectors. Specifically, it means we would estimate 428×428 elements of the knowledge diffusion matrix.

5.1 Calibration of Parameters

We assume that the distribution of idiosyncratic fixed costs of research, $H(\zeta)$, is gamma with mean one and variance σ_ζ^2 .⁴⁵ The shocks to individual firm's innovation and imitation are also drawn from a gamma distribution $G(\varepsilon)$ with mean zero and variance σ_ε^2 . The set of parameters in the model to be calibrated includes elements of the intersectoral knowledge diffusion matrix, A , and other parameters $\{\beta, \alpha, \theta, \sigma, \eta, F, \sigma_\zeta, \sigma_\varepsilon\}$.

intersectoral knowledge diffusion matrix A We proxy the knowledge linkage by the fraction of citations made to sector j by sector i (knowledge flow from sector j to i). Since sectors with more patents tend to be cited more frequently, we handle this by normalizing the citation percentage by the relative importance of sector j , measured by the share of citations received by j in total citations, citationshare^j .⁴⁶ Formally,

$$\tilde{A}^{i \leftarrow j} = \frac{\text{no. of citations from } i \text{ to } j / \text{total citations made by } i}{\text{citationshare}^j}. \quad (29)$$

Figure VII shows a contour graph of the knowledge diffusion matrix, $(\log)\tilde{A}^{i \leftarrow j}$ for these 42 sectors. The darkest area on the diagonal reflects the fact that a large proportion of citations go to patents in the same sector. This is not particularly surprising given that sectors in this case are not highly disaggregated; however, most sectors also allocate a significant amount of citations to patents from other sectors, reflecting the importance of cross-sector knowledge spillovers. We normalize the knowledge diffusion matrix by a scale parameter, A_0 , such that $A^{i \leftarrow j} = A_0 \times \tilde{A}^{i \leftarrow j}$. This parameter governs the average innovation productivity over all sector pairs.

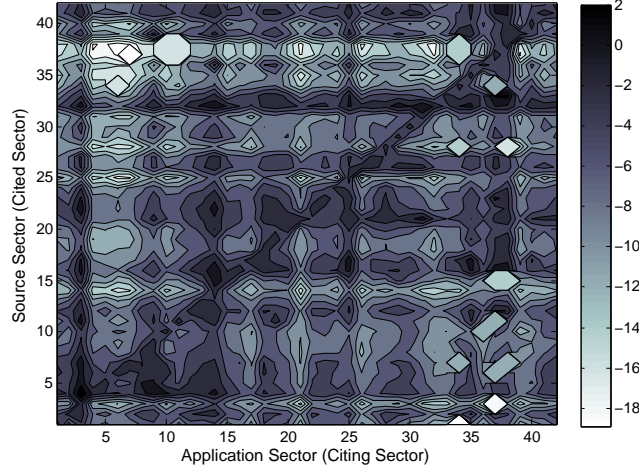
Other Parameters Total labor force, L , is normalized to 100 and we choose the following standard numbers: the (gross) growth rate of real output $g = 1.02$, the interest rate $r = 0.05$ and $\beta = 0.99$. The average (gross) growth rate of patents, γ , equals 1.11 in the data (for the period 1976-2006). (23) and (28) then imply that the household's risk aversion parameter $\eta = 3$, and the elasticity of substitution between differentiated goods $\sigma = 6$, which lies within the range of estimates of elasticities of substitution provided in Anderson and Van Wincoop (2004) and Broda and Weinstein (2006).⁴⁷ Unfortunately, there is no direct information to pin down the rest

⁴⁵The scale and shape parameters of this gamma distribution are σ_ζ^2 and $1/\sigma_\zeta^2$, respectively. The theory in Kesten (1973) works with many types of distributions of shocks as long as shocks are i.i.d over time. We choose gamma distribution because the random number generator in Matlab program behaves well under this distribution for our sample size.

⁴⁶It is important to note that this measure is different from other technology closeness measures proposed in Jaffe (1986) and Bloom et al. (2010). These previous papers study the bilateral distance between any two technologies which is independent from the direction of knowledge flows. In our paper, the knowledge diffusion matrix is asymmetric across sector-pairs, i.e. $\tilde{A}^{i \leftarrow j} \neq \tilde{A}^{j \leftarrow i}$.

⁴⁷Using detailed imports data, Broda and Weinstein (2006) estimate the elasticities of substitution between differ-

Figure VII: Contour Graph of Knowledge Diffusion Across Sectors



Note: The figure represents the knowledge diffusion matrix constructed from the NBER Patent Citation data for 42 technological sectors. A darker color implies that the sector is cited by another at a higher rate than a sector-pair with a lighter color.

of the model parameters. Therefore, we combine the Generalized Method of Moments (GMM) and Simulated Method of Moments (SMM) and use firm patenting information to estimate these parameters.

First, we use GMM to back out the parameters that enter the aggregate equilibrium conditions (13), (21), (22), and (26): $A_0, \alpha, \theta, \{v^i\}_i$. Specifically, we adopt the continuously updating GMM, where the optimal weighting matrix is estimated simultaneously with the parameter values. Pooling the patent data for the period 1976–2006, we observe firm’s patenting behavior in these 42 sectors. Based on this information, we calculate the 30-year average of the relative patent stock across sectors, $\{n^i/n^j\}_{ij}$, the average fraction of firms in each sector, $\{M^i/M\}_i$, and the fraction of patents in sector i owned by firms that have previously innovated in j , $\{\tau^{i \leftarrow j}\}_{i,j}$. In addition, we obtain the ratio between the number of firms and total population, M/L , from Axtell (2011).⁴⁸

Now define vector $\vartheta \equiv \{A_0, \alpha, \theta, \{v^i\}_i\}$ and $G_t(\vartheta)$ is the vector of differences between real data moments and equilibrium model moments. $T = 30$ is the total number of periods. The moments that we attempt to match are Equations (13), (21), (22) and (26). Thus, there are 87 equations

entiated goods for sectors at various disaggregated level. The average of the elasticities of substitution is 6.8 among 3-digit SITC goods during 1972-1988 and 4 during 1990-2001. Anderson and van Wincoop (2004) review the previous studies and conclude that the elasticity of substitution is likely to be in the range of five to ten.

⁴⁸there are 5.07 million firms in the U.S. and the total population is 249 million in 1990.

and 46 unknowns in our estimation. Our estimator minimizes

$$\hat{\vartheta} = \arg \min_{\vartheta} \left[\frac{1}{T} \sum_{t=1}^T G_t(\vartheta) \right] \left[\frac{1}{T} \sum_{t=1}^T G_t(\vartheta)' G_t(\vartheta) \right]^{-1} \left[\frac{1}{T} \sum_{t=1}^T G_t(\vartheta) \right]'$$

We then use SMM to estimate the remaining parameters, F , σ_ε and σ_ζ .⁴⁹ Our target moments are: the 30-year average of the mean number of sectors per firm $\bar{S} = 2.61$; the fraction of firms in each sectors $\{M^i/M\}$; and the shape parameter of the Pareto distribution of patent stock across firms, $\mu = 1.89$. \bar{S} helps to pin down the average fixed cost F . Information on the distribution of M^i/M helps to discipline σ_ζ , because when shocks to sectoral fixed costs are more volatile, firms' sectoral selection becomes more random, and thus the distribution of the number of firms across sectors becomes more even. Similarly, more volatile innovation shocks increases the heterogeneity of firms' patent stock; thus, information of μ helps to identify σ_ε .

There are $M=33,000$ firms in the simulation, which is the maximum number of active firms within one year in the patent data. Every period, based on (18), $\bar{S} \times M$ firm-sector pairs are selected to be actively conducting R&D. The active firm-sector pairs follow firm dynamics in (24). F is estimated using the average realized fixed costs of these active firms in the simulation. For any pair of σ_ε and σ_ζ , the simulation starts from the firm patent stock distribution in 1997.⁵⁰ We then repeat the following steps for $2T = 100$ periods.

1. At period t , define and calculate the expected value gain of each firm f in sector i if firm f chooses to innovate in sector i (after the realization of shocks to its fixed costs of R&D $\zeta_{f,t}^i$ for every i) as

$$score_{f,t}^i = \sum_{j=1}^K \omega^{i \leftarrow j} \frac{(z_{f,t}^j + \theta \bar{z}_t^j)}{n_t^j} \times (\zeta_{f,t}^i)^{-1},$$

where $\{\omega^{i \leftarrow j}\}_{i,j}$ are calculated according to (14), using parameters estimated previously.

2. Select the cutoff value $F_{\max,t}$ such that only $\bar{S} \times M$ elements among all $\{score_f^i\}_{f,i}$ are greater than $F_{\max,t}$ in period t . Firm dynamics follow (24) if $score_f^i > F_{\max,t}$; otherwise, the firm is idle in sector i for period t . F_t is estimated using the average fixed costs faced by active firms at time t . $F(\sigma_\varepsilon, \sigma_\zeta)$ is then estimated by taking the average of such F_t s in the last T periods.

In the end we choose the pair of σ_ε and σ_ζ that minimizes the quadratic distance between the

⁴⁹The simulation process is extremely time consuming when keeping track of a large number of firms ($N = 33,000$) and their innovation outcome in 42 sectors. We estimate σ_ε and σ_ζ using SMM because they do not enter the aggregate equilibrium conditions analytically. Moreover, in order to implement the equilibrium firm dynamics governed by (24) and firm's entry decision in (18), we need to know other model parameters estimated previously.

⁵⁰We choose 1997 because the number of patenting firms is the largest. We also assume that the firm over population ratio does not change over time.

Table II: Parameter Values

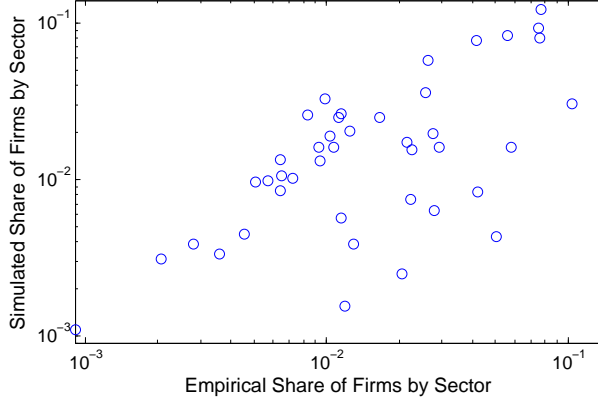
β	σ	η	α	θ	A	F	σ_ε	σ_ζ
0.99	6	3	0.44	0.00376	0.0038	0.001	1	2

simulated moments (from the last T period) and the empirical moments.

The calibrated parameter values are reported in Table II. Most notably, $\alpha = 0.44$, which implies a substantial input from researchers in the knowledge creation process. The imitation efficiency parameter $\theta = 0.0038$ suggests that private knowledge previously accumulated is significantly more important than public knowledge for innovating firms.

5.2 Goodness of Fit at Sectoral Level

Figure VIII: Empirical and Model Generated Number of Firms Across Sectors



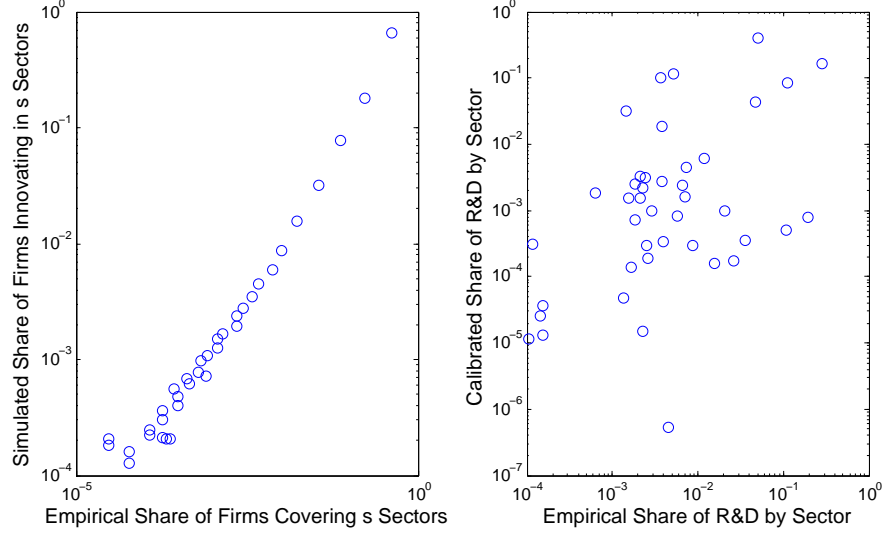
We show the goodness of fit of our model by comparing the model's targeted and untargeted moments with their empirical counterparts. In Figure VIII we plot the cross-sector observations of the model generated number of firms and empirical number of firms across sectors (correlation=0.62). The shape parameter of the Pareto distribution of firms' total patent stock is 1.91 in the simulation, close to $\mu = 1.87$ in the data.

We also investigate our model's prediction about some of the moments that we did not directly target. The left panel in Figure IX shows that the simulated number of firms that patent in s number of sectors $\{N^s\}$ in the model is highly correlated with that in the data (correlation=0.99). The right panel compares the estimated share of R&D expenditure across sectors using our model and real sectoral R&D expenditures in the data.⁵¹ Again, we find a value of 0.64 for the correlation

⁵¹Equation (25) implies that the R&D expenditure of sector i is proportional to the market value of sector i 's knowledge v^i .

between these two variables. Additionally, the model also replicates another set of untargeted moments, the share of knowledge in sector j that are utilized in the R&D of sector i $\{\tau^{i \leftarrow j}\}$ fairly well, the correlation between the simulated moments and their empirical counterparts is as high as 0.61.

Figure IX: Untarged Moments: Cross-sector Distribution of Firms and R&D



5.3 Simulated Firm-level Observations

When presenting the simulation results, we divide the large number of simulated firms into 20 bins (similar to Section 2) and present the observations based on an average firms in each bin.⁵² Results are shown in Figure X–XIII.

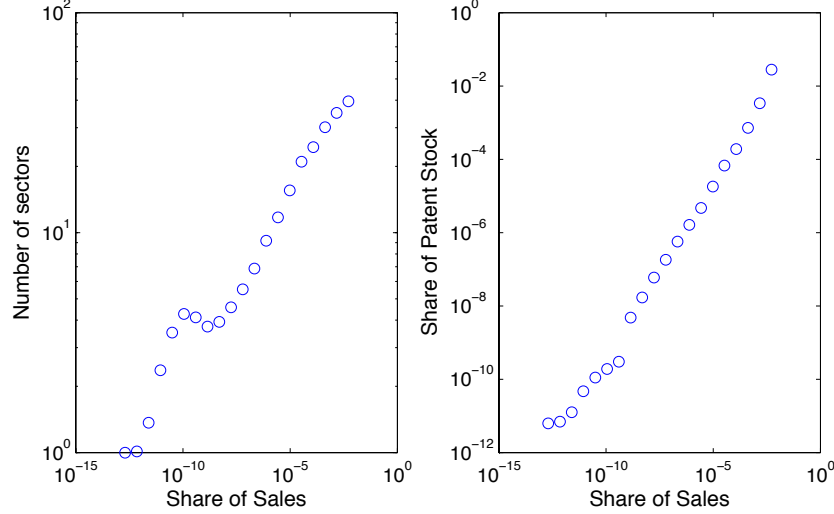
Firm Size, Patent Stock and Multi-Technology Patenting In the data, we observe significantly positive relationship between firm size and firm patent performance (patent stock and multi-technology patenting) (Table I). Figure X shows that this is captured in the model as well (firm size in the model is measured by it total sales $\sum_{i=1}^K \frac{z_{f,t}^i}{n_t^i} s^i PY$).⁵³ We find that the correlation

⁵²Bins are evenly divided according to the log-scaled X-axis variable. Therefore there may be uneven number of observations in all bins.

⁵³Since all varieties in the same sector are sold at the same amount and charge the same price in the model, firm's knowledge capital in a specific sector—measured by the number of its past innovations—is proportional to the firm's sales in that sector. Therefore, a one-sector model would naturally generate a positive and perfect correlation between firm size and firm's knowledge capital. This is, however, not as straightforward in a multi-sector model because the total knowledge capital can differ across sectors. Depending on the sectors in which a firm innovates, a large number of past innovations may not necessarily imply large sales as the firm's market share can be small in these sectors (a firm's total sales in the model are defined as $\sum_{i=1}^K \frac{z_{f,t}^i}{n_t^i} s^i PY$). Thus, we use simulation to test the prediction of our model.

between the simulated number of varieties and sales equals 0.92 and the correlation between the number of sectors in which a firm has one or more patents and firm's sales equals 0.71.

Figure X: Larger firms have more patents and innovate in more patent classes

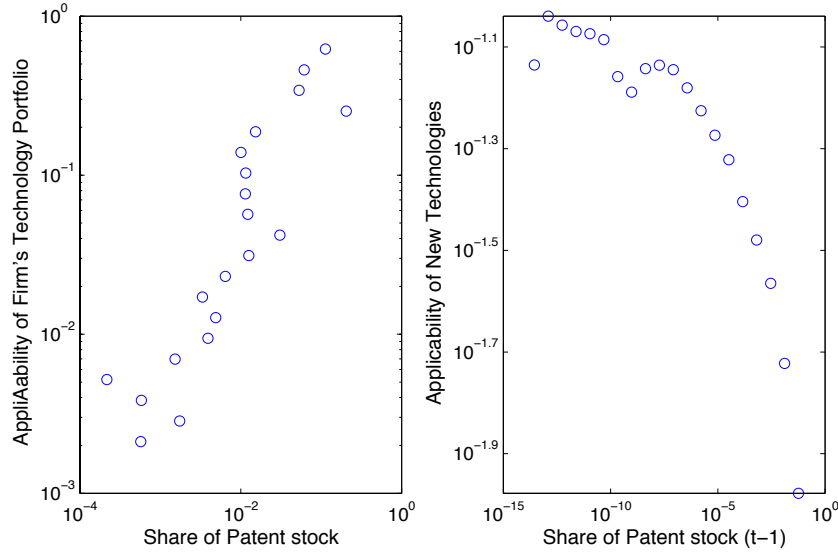


Knowledge Capital and Technology Applicability In line with the empirical observations in Figure V, our simulation shows that the model predicts a similar relationship between a firm's knowledge capital (measured by the number of blueprints of the firm, $\sum_i z_f^i$) and its position in the technology space. Firms with larger knowledge capital tend to concentrate more in sectors with high knowledge applicability, because they are better at internalizing knowledge spillovers across sectors and thus have better incentive to innovate in central sectors.

As described in the left panel of Figure XI, firm's technology applicability (defined in the same way as in the Section 2—weighted average applicability of its technologies using authority weight calculated for these 42 sectors in the simulation) increases with its total number of innovations. However, since firms tend to enter multiple sectors sequentially, the knowledge applicability of the recent sectors that a firm enter is negatively related to firm's knowledge capital (as shown in the right panel of Figure XI).

Firm Innovation Rate and Initial Technology Applicability Figure XII plots the simulated firm innovation rate, $\frac{\sum_i \Delta z_{f,1}^i}{\sum_i z_{f,0}^i}$, against its initial applicability of its technology portfolio. It is evident that the model replicates the empirical observation that firms' innovation rates are positively related to the applicability of their initial technological position. A more central positioning in the technology space opens more potential routes for a firm to expand across sectors, thus boosting the firm's extensive growth (innovation in new sectors). It also allows the firm to apply knowledge from many related sectors to innovating in the existing sectors, driving up the firm's intensive growth.

Figure XI: Firms with more patents concentrate more in highly applicable technologies. Yet, their recent patents are in the less applicable ones.



Pareto Tail Distribution of Firm Sizes Our simulation shows that the distribution of firm sizes (measured by the average share of patent stock, $\frac{1}{K} \sum_i \frac{z_f^i}{n^i}$, in the model) is indeed well approximated by a Pareto distribution. Figure XIII depicts the stationary firm size distribution by plotting the $\log(rank)$ against $\log(size)$ for large firms. It shows that not only the shape parameter for the simulated distribution is close to the estimated parameter from the data. The relationship between $\log(rank)$ and $\log(size)$ is well approximated by a straight line (except for the few observations of very large firms), indicating a Pareto distribution (as in Gabaix, 1999 and Acemoglu and Cao, 2010).

6 Final Remarks

Technological advances are complementary and sequential; interconnections between them are, however, highly heterogeneous. Our goal is to forge a link between observations of firm innovation in multiple technologies and theories of aggregate technological progress. We provide a theoretical framework which builds on micro-level observations and helps to elucidate how innovating firms choose to position themselves in the technology space and allocate their R&D investment. We have attempted to demonstrate that our model can replicate key firm-level facts; as such, the resulting aggregate model is likely to provide a more credible tool for policy analysis.

Our study has important implications for economic growth and R&D policies. First, government policies directed at stimulating innovation in certain technologies need to be based on better

Figure XII: Firms innovate faster if their initial technologies are more applicable

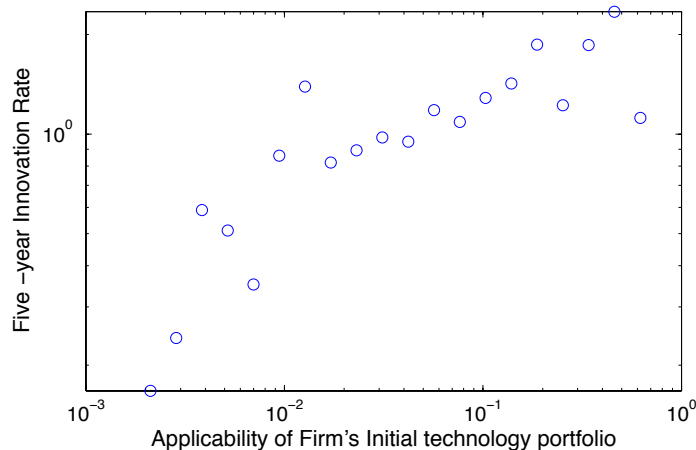
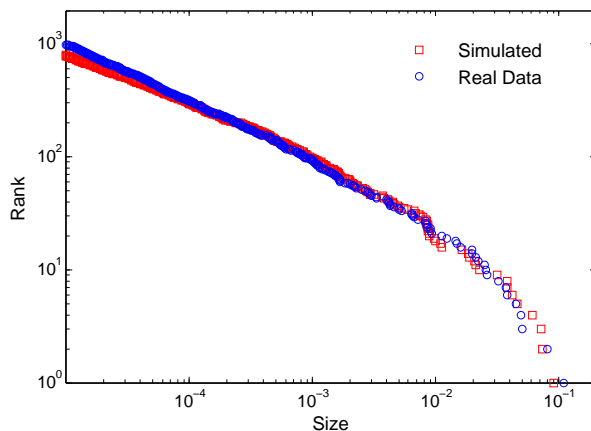


Figure XIII: Pareto Distribution of Firm Sizes



understanding of the intersectoral knowledge linkages. Heterogenous sectoral knowledge spillovers suggest that industrial or R&D policies that favor highly applicable sectors may boost growth.⁵⁴ Second, policies that lower barriers to diversification help to reinforce the effect of industrial policies, as it can be challenging to shift to more advanced industries given the fixed cost of learning and adapting technology in new sectors, and more diversification encourages spillovers between different technologies. Third, competition policies that encourage joint R&D ventures in highly related sectors can benefit growth, because firms are better able to internalize knowledge spillovers.⁵⁵

One direction for future research is to provide a better understanding of the pattern of sequential

⁵⁴In fact, in a cross-country study, Cai and Li (2012) show that countries which specialize more in applicable technologies tend to grow faster.

⁵⁵A successful example is China. Over the past two decades, China has significantly shifted its industrial structure from specializing in exporting low or medium knowledge applicable (e.g. “Textile mill products” and “Food and kindred products”) to exporting proportionally more highly applicable products (e.g. “Electronic components and communications equipment” and “Office computing and accounting machines”). The Chinese government has adopted a set of policies promoting structural transformation.

sectoral entry using both firm innovation and production data. Our empirical findings suggest that firms follow a general pattern when they move through the technology space: firms tend to start from highly applicable central technologies and gradually expand to related technologies towards the fringe of the technology space. This sequential sectoral entry has the potential to explain observations at firm and sector levels other than those shown in the current paper. For example, cross-sector knowledge linkages could potentially help understand the non-random sectors co-production phenomenon documented by Bernard, Redding and Schott (2010), which finds that some pairs of sectors (e.g., fabricated metal and industrial machinery) are systematically more likely to be produced by the same firms than other sector pairs. Our analysis suggests that the source of these differences could arise from heterogeneity in cross-sector knowledge spillovers. In fact, we find that the likelihood of innovating in sector i conditional on having already innovated in sector j —measured by the average percentage of firms in j that also have patents in i —positively correlates with our measure of intersectoral knowledge linkages, $\log(A^{ij})_{i \neq j}$, at 56%. However, this correlation only shows the impact of knowledge in an upstream sector on innovation in an immediate downstream sector, hence presents an incomplete story of sectors co-R&D. For instance, a firm may innovate in both i and j , because j provides substantial knowledge spillovers to a third sector k —which contributes significantly to innovations in i . Since it is not the focus of the paper, we leave for future research a greater in-depth analysis on this topic.

Another unexplored prediction of our model is that a firm’s market value should increase with the applicability of its technology portfolio.⁵⁶ Empirical investigation of these predictions could also be interesting for future research.

Our analysis of multi-sector firm innovation in the presence of barriers to diversity also prompts the questions: What are quantitative implications of these barriers? What are the counterfactuals? What are the appropriate government policies to mitigate the potential inefficiencies in a competitive economy? We leave these questions for future work in a separate paper.

⁵⁶Hall, Thoma and Torrisi (2007) find that Tobin’s q is significantly positively associated with a firm’s R&D and patent stock, and modestly increases with the quality of patents (measured by forward citations).

References

- [1] Acemoglu, D. and D. Cao. 2010. “Innovation by Entrants and Incumbents”, *NBER Working Paper* No. 16411.
- [2] Acemoglu, D, Carvalho, V., Ozdaglar and A. Tahbaz-Salehi. 2012. “The Network Origins of Aggregate Fluctuations” *Econometrica*, forthcoming.
- [3] Aghion P. and P. Howitt. 1992. “A Model of Growth Through Creative Destruction”. *Econometrica*, 60(2): 323-351.
- [4] Akcigit U. and W. Kerr. 2010. “Growth Through Heterogeneous Innovations”, *NBER working paper* No. 16443.
- [5] Akcigit U., D. Hanley and N. Serrano-Velarde. 2012. “Back to Basics: Basic Research Spillovers, Innovation Policy and Growth”, University of Pennsylvania, working paper.
- [6] Anderson J. and E. Van Wincoop. 2004. “Trade Costs”. *Journal of Economic Literature*, 42(3): 691-751.
- [7] Antras, P., D. Chor, T. Fally and R. Hillberry. 2012. “Measuring the Upstreamness of Production and Trade Flows”. *American Economic Review*, 102(3):412-416.
- [8] Atkeson, A. and A. Burstein. 2010 “Innovation, Firm Dynamics and International Trade”, *Journal of Political Economy*, 118(3): 433-484.
- [9] Axtell, R.. 2001 “Zipf Distribution of U.S. Firm Sizes”, *Science*, 293(5536): 1818-1820.
- [10] Balasubramanian, N. and J. Sivadasan. 2011. “What Happens When Firms Patent? New Evidence from U.S. Economic Census Data”. *Review of Economics and Statistics*, 93(1): 126-146.
- [11] Barseghyan, L. and R. DiCecio. 2009. “Entry Costs, Industry Structure, and Cross-Country Income and TFP Differences”. *Journal of Economic Theory*, 146(5): 1828-1851.
- [12] Basu, S., 1995. “Intermediate goods and business cycles: implications for productivity and welfare”. *American Economic Review*, 85:512- 531.
- [13] Belenzon, S. and M. Schankerman. 2010. “Spreading the Word: Geography, Policy and University Knowledge Diffusion”. *CEPR Discussion Paper* No. 800
- [14] Bernard, A. B., S. Redding, and P. K. Schott. 2011. “Multi-Product Firms and Trade Liberalization”. *Quarterly Journal of Economics*, 126(3): 1271-1318.
- [15] Bernard, A., S. J. Redding and P. K. Schott. 2010. “Multi-Product Firms and Product Switching”. *American Economic Review*, 100(1): 70-97
- [16] Bernard, A., S. J. Redding and P. K. Schott. 2009. “Product and Productivity”. *NBER working paper* No.11575.
- [17] Bernstein and Nadiri. 1988. “Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries”. *American Economic Review*, 78(2): 429-439.
- [18] Bloom, N., M. Schankerman and J. V. Reenen. 2010. “Identifying Technology Spillovers and Product Market Rivalry”. working paper.

- [19] Boedo, H. and T. Mukoyama. 2010. "Evaluating the Effects of Entry Regulations and Firing Costs on International Income Differences". University of Virginia, working paper.
- [20] Bresnahan, T., and M. Trajtenberg. 1995. "General Purpose Technologies: 'Engines of Growth'?", *Journal of Econometrics*, 65(1): 83-108.
- [21] Brada, C. and D.E.Weinstein. 2006. "Globalization and The Gains From Variety", *Quarterly Journal of Economics*, 121(2): 541-585.
- [22] Buera, F., J. Kaboski and Y. Shin. 2011. "Finance and Development: A Tale of Two Sectors", *American Economic Review*, 101(5):1964-2002.
- [23] Cai, J. 2010. "Knowledge Spillovers and Firm Size Heterogeneity". University of New South Wales, working paper.
- [24] Cai, J. and N. Li. 2012. "Intersectoral Knowledge Linkages, Technology Compositions and Growth". Ohio State University, working paper.
- [25] Carvalho, V. 2010. "Aggregate Fluctuations and the Net Work Structure of Intersectoral Trade". University of Chicago, mimeo.
- [26] Conley, T. and B. Dupor. 2003. "A Spatial Analysis of Sectoral Complementarity ". *The Journal of Political Economy*, 111(2): 311-352.
- [27] David, P. 1990. "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox". *American Economic Review*, Vol. 80, No. 2, pp. 355-361.
- [28] Gabaix, X. 2009. "Power Laws in Economics and Finance". *Annual Review of Economics*, Vol. 1, pp. 255-294.
- [29] Grossman, G. and E. Helpman. 1991a. "Innovation and Growth in the Global Economy". Cambridge, MA: MIT Press.
- [30] Grossman, G. and E. Helpman. 1991b. "Quality Ladders in the Theory of Growth". *Review of Economic Studies*, 68:43-61.
- [31] Hall, B, A. Jaffe, and M. Trajtenberg. 2001. "The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools". NBER Working Paper 8498.
- [32] Hausmann, R., J. Hwang and D. Rodrik. 2005. "What You Export Matters". *Journal of Economic Growth*, 12(1):1-25.
- [33] Helpman, E. (ed.). 1998. "General Purpose Technologies and Economic Growth". Cambridge and London: MIT Press.
- [34] Hidalgo, C.A., B. Klinger, A.-L. Barabasi and R. Hausman. 2007. "The Product Space Conditions the Development of Nations". *Science*, 317:482-487.
- [35] Hirschman, A. O. 1958. "The Strategy of Economic Development". Yale University Press.
- [36] Hopenhayn, H. 1992. "Entry, Exit and Firm Dynamics in Long Run Equilibrium". *Econometrica*, 60(5): 1127-50.
- [37] Horvath, M. 1998. "Cyclicalities and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks". *Review of Economic Dynamics*, Vol. 1, pp. 781-808.

- [38] Ilyina, A. and Samaniego, R. M. 2011. "Technology and Financial Development". *Journal of Money, Credit and Banking* 43(5): 899-921.
- [39] Jovanovic B. and P. L. Rousseau. 2005. "General Purpose Technologies". In Philippe Aghion and Steven Durlauf (Eds.), *Handbook of Economic Growth*, 1: 1181-1224.
- [40] Jaffe, Adam. 2012. "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value". *American Economic Review*, 76(5): 984-1001.
- [41] Jaffe, A. 1988. "Demand and Supply Influences in R&D Intensity and Productivity Growth", *Review of Economics and Statistics*, 120 (3): 431-437.
- [42] Jones, C. I. 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development". *American Economic Journal: Macroeconomics*, 3(2): 1-28.
- [43] Jones, C. I. 1995. "R&D-Based Models of Economic Growth". *Journal of Political Economy*, 103(4):759-784.
- [44] Jones, C. I. 1999. "Growth: With or without scale effects?" *Papers and Proceedings of the One Hundred Eleventh Annual Meeting of the American Economic Association*, 89: 139-144.
- [45] Kesten, H. 1973. "Random Difference Equations and Renewal Theory for Products of Random Matrices". *Acta Mathematica*, 131: 207-248.
- [46] Kleinberg, R. 1999. "Authoritative Sources in a Hyperlinked Environment" *Journal of Association for Computing Machinery*, 46(5): 604-632.
- [47] Klenow, P. 1996. "Industry Innovation: Where and why?" *Carnegie-Rochester Conference Series on Public Policy*, 44: 125-150.
- [48] Klette, T. J. and S. Kortum. 2004. "Innovating Firms and Aggregate Innovation". *Journal of Political Economy*, 112: 986-1018.
- [49] Landes, D. 1969. "The unbound Prometheus". Cambridge University Press.
- [50] Levin, R.C., Cohen, W.M., and Mowery, D.C. 1985. "R&D Appropriability, Opportunity, and Market Structure: New Evidence on Some Schumpeterian Hypotheses", *American Economic Review*, 75: 20-24.
- [51] Leontief, W. 1936. "Quantitative Input and Output Relations in the Economic System of the United States". *Review of Economics and Statistics*, Vol. 18, No. 3, pp.105-125.
- [52] Lucas, R.E. 1981. "Understanding Business Cycles". In *Studies in Business Cycle Theory*. Cambridge, MA: MIT Press.
- [53] Luttmer, E. G. 2012. "Technology Diffusion and Growth". *Journal of Economic Theory*, 147: 602-622.
- [54] Luttmer, E. G. 2007. "Selection, Growth, and the Size Distribution of Firms". *Quarterly Journal of Economics*, Vol. 122, No. 3, pp.1103-1144.
- [55] Mayer, Melitz and Ottaviano. 2011. "Market Size, Competition, and the Product Mix of Exporters," *NBER Working Papers* No. 16959.

- [56] Nelson, R. and S. Winter. 1977. "In Search of Useful Theory of Innovation". *Research Policy*, 6: 36-76.
- [57] Ngai, R. and R. Samiengo. 2011. "Accounting for Research and Productivity Growth Across Industries". *Review of Economic Dynamics*, 14(3), 475-495.
- [58] Pakes, A. and M. Schankerman. 1984. "An Exploration into the Determinants of Research Intensity", Zvi Griliches, (ed.), *R&D Patents, and Productivity*, University of Chicago Press, 209-232.
- [59] Palacios-Huerta, I. and O. Volij. 2004. "The Measurement of Intellectual Influence". *Econometrica*, 72(3): 963-977.
- [60] Pinski, G., and F. Narin. 1976. "Citation Influence of Journal Aggregates of Scientific Publications: Theory, with Applications to the Literature of Physics". *Information Processing and Management*, 12: 297-312.
- [61] Rajan, R. G. and L. Zingales. 1998. "Financial Dependence and Growth". *American Economic Review*, 88(3): 559-586.
- [62] Romer, P. M. 1990. "Endogenous Technological Change". *Journal of Political Economy*, 98: 71-102.
- [63] Rosenberg, N. 1982. "Inside the black box: Technology and economics". Cambridge University Press.
- [64] Rossi-Hansberg E. and Wright M. 2007. "Establishment Size Dynamics in the Aggregate Economy" *American Economic Review*, 97(5): 1639-1666.
- [65] Segerstrom, P. S. , T.C. A. Anant, and E. Dinopoloulos. 1990. "A Schumpeterian Model of the Product Life Cycle". *American Economic Reviews*, 80(5):1077-?1091.
- [66] Wieser, R. 2005. "Research and Development Productivity and Spillovers: Empirical Evidence at the Firm Level". *Journal of Economic Survey*, 19(4) : 587-621.

A Data Appendix

A.1 Data Sources

Firm Patenting and Patent Citations We use patent applications in the 2006 edition of the NBER Patent Citation Data (see Hall, Jaffe and Trajtenberg, 2001 for details) to characterize firms’ innovation activities and their citations to trace the direction and intensity of knowledge flows and to construct indices of knowledge linkages among sectors. The data provides detailed information of every patent granted by the United States Patent and Trade Office (USPTO) and their citations from 1976 to 2006. We summarize each firm’s patent stock in each disaggregated technological class (intensive margin of innovation) and the number of categories (extensive margin of innovation) for each year.

Each patent corresponds to one of the 428 3-digit United States Patent Classification System (USPCS) technological classes and also one of more than 800 7-digit International Patent Classification (IPC) classes. Figure II presents the intersectoral network corresponding to patent citation share matrix for 428 technological category. We mostly report the results based on USPCS codes, but we check for robustness using the IPC classes. We also present some evidence based on industrial sector classification, as the model is estimated based on this categorization. To translate the data into the industrial classifications, we use the 2005 edition of the concordance table provided by the USPTO to map USPCS into SIC72 (Standard Industrial Classification in 1972) codes, which constructs 42 industrial sectors.⁵⁷ We summarize citations made to patents that belong to the same technological class and use the cross-sector citations to form the intersectoral knowledge spillover network.

Firm R&D and Accounting Data Information on firm sizes (i.e. sales or employment) and firm’s R&D expenditure is from the U.S. Compustat database. Firm-level R&D intensity is defined as R&D expenditure divided by sales. The industry measure of R&D intensity is the median firm value. For robustness check, we also investigated the relationship between sectoral R&D intensity and sectoral knowledge applicability using the average firm instead: again, they are significantly positively related.

To obtain information of firm sizes (i.e. number of employees, sales), we use the NBER’s mapping between the Compustat data and the patent data between 1970 to 2006 and keep only patenting firms.

⁵⁷The patents are classified according to either the intrinsic nature of the invention or the function for which the invention is used or applied. It is inherently difficult to allocate the technological category to economically relevant industries in a differentiation finer than 42 sectors, even with detailed firm level information. First, most of the patents are issued by multi-product firms that are present in multiple SIC-4 industries. Second, in the best scenario, one only has industry information about the origin of the patents but not the industry to which the patent is actually applied.

A.2 Alternative Measures of Sectoral Knowledge Applicability

Technology Distance Based on the patent citation network, we also construct a pairwise knowledge distance measure to facilitate our studies. Define a $K \times K$ distance matrix D , where the $\{i, j\}^{th}$ element, $D^{ij} = d$ if $(C^\wedge d)^{ij} > 0$ and $(C^\wedge (d-1))^{ij} = 0$. $C^\wedge d$ denotes the d^{th} power of the matrix C . D^{ij} is the shortest path distance between the nodes i and j . If $(C^\wedge d)^{ij} > 0$, there is at least one indirect route via other $d-1$ nodes between nodes i and j . If $(C^\wedge d)^{ij} > 0$, that means there exists at least one d -step route between i and j . If $(C^\wedge (d-1))^{ij} = 0$ is also true, then d is the shortest path distance between i and j .

The mean of D 's i^{th} column is the average distance between technology i and all other technologies. We find the average distance to other technologies is negatively correlated (-0.58) with the authority weight, since higher authority weight products are located closer to the center of the network, which are connected to more other technologies. The negative correlation is not perfect, because the average distance ignores the volume of knowledge flow between sectors and the importance of related sectors. Nevertheless, the distance measure helps to understand the relatedness between technologies (or sectors).

Upstreamness Using U.S. Input-Output (I-O) Table, Antras et al. (2012) construct an industry-level measure of relative production line position. The intersectoral knowledge diffusion network can be interpreted as knowledge input-output matrix across different technologies. Thus, following a similar method as in Antras et al. (2012), we construct a measure of knowledge upstreamness as follows.

Denote the upstreamness measure of technology category i as U_i . $C_{j \rightarrow i}$ is the number of citations made by category- j patents to previous patents in category i . $C_{\rightarrow i}$ is the total number of citations received by patents in category i . $citationshare_i$ is the share of citations made to patents in category i among all citations between 1976–2006. $\frac{citationshare_i}{\max(citationshare_i)} \frac{C_{j \rightarrow i}}{C_{\rightarrow i}}$ indicates the weighted share of i 's knowledge used as intermediate inputs by j (similar to the share of sector i 's output purchased by sector j as in Antras et al. 2012). The weight $\frac{citationshare_i}{\max(citationshare_i)}$ is chosen such that the most cited category is assumed to be 100% used as knowledge intermediate input, while if a sector does not attract any citation, its knowledge is assumed to be only used in its own sector.⁵⁸

$$U_i = 1 + \sum_{j=1}^K \frac{citationshare_i}{\max(citationshare_i)} \frac{C_{j \rightarrow i}}{C_{\rightarrow i}} U_j$$

In-degree In the patent citation network, in-degree of a technology class i is the total number of in-ward citation linkages from other classes to a certain technology class. Compared to authority weight, in-degree ignores the importance of the citing class and the citation volume of each in-ward

⁵⁸The upstream sectors in the knowledge I-O table are different from the upstream sectors in the production I-O table. In the knowledge I-O table, the most upstream sectors are the frontier and general purpose technologies that are widely adopted by many other sectors, while the most downstream sectors are the mature and specific purpose technologies. In the production I-O table, the most upstream industries are related to metal materials, while the most downstream industries are the final consumption goods, such as automobile and footwear.

linkage. Still, it provides a rough measure of class i 's knowledge contribution to the entire network.

Table A.1 presents the bilateral correlation between different measures of sectoral knowledge applicability. It is evident that they are all significantly and positively correlated. In fact, all the empirical observations we documented in Section 2 are robust to using different measures.

Table A.1: Correlation between Different Measures of Sectoral Knowledge Applicability

Variables (in log-scale)	aw	1/avg. distance	in-degree	upstreamness
aw	1			
1/avg. distance	0.579	1		
in-degree	0.786	0.375	1	
upstreamness	0.620	0.632	0.500	1

A.3 Firm-level Observations: Regression Results

Observation 2 and 3: In Figure V, the average applicability of the firms' patent portfolio increases with their total patent stock and numbers of categories or technology scope. In the following regressions, we study carefully how firms allocate their innovative efforts across sectors and present the results in Table A.2. The regressions are based on the following specifications:

$$S_{f,t}^i/S_{f,t} = \beta_1 S_{f,t} + \beta_2 \ln aw^i + \beta_3 S_{f,t} \times \ln aw^i + d_f + d_t$$

$$\Delta S_{f,t}^i/\Delta S_{f,t} = \beta_1 S_{f,t} + \beta_2 \ln aw^i + \beta_3 S_{f,t} \times \ln aw^i + d_f + d_t$$

$S_{f,t}^i/S_{f,t}$ is the share of patent stock allocated to category i by firm f . $\Delta S_{f,t}^i/\Delta S_{f,t}$ is the share of new patents allocated to category i by firm f . $\beta_1 < 0$ implies that a larger firm spreads a smaller share of total R&D effort in an average technology category. $\beta_2 < 0$ implies an average firm allocates less R&D effort to a highly applicable sector than other sectors. $\beta_3 > 0$ implies that a larger firm allocates greater share of patents in the highly applicable center sectors than smaller firms. We confirm the robustness of these relationships at different levels of disaggregation: 3-digit SIC industry level (42 sectors), US patent categories (Nclass, 428 sectors) and the International Patent Classification (IPC) classification (977 sectors).

Further, in Table A.3 we study how firms expand across different technological categories given the heterogenous technology applicability. To investigate the innovation patterns over time, we run the following two fixed effect regressions, controlling for firm fixed effects in each case.

$$\begin{aligned} \ln aw_{f,t}^i &= \beta_0 + \beta_1 \ln S_{f,t} + \beta_2 new_{f,t}^i + d_f + d_t + \varepsilon_{if,t} \\ \ln aw_{f,t}^i &= \beta'_0 + \beta'_1 \ln S_{f,t} + \beta'_2 new_{f,t}^i \cdot \ln S_{f,t} + d_f + d_t + \varepsilon_{if,t} \end{aligned}$$

where $S_{f,t}$ is firm f 's patent stocks over all sectors, $new_{f,t}^i$ is a dummy variable equal to one if firm f is a new entrant in sector i at time t and $aw_{f,t}^i$ is the authority weight of sector i in which firm f

Table A.2: R&D Allocation and Firm Size

Dependent Variables Independent Variable	3-digit SIC		NClass Categories		IPC Categories	
	$\log \frac{S_{f,t}^i}{S_{f,t}}$	$\log \frac{\Delta S_{f,t}^i}{\Delta S_{f,t}}$	$\log \frac{S_{f,t}^i}{S_{f,t}}$	$\log \frac{\Delta S_{f,t}^i}{\Delta S_{f,t}}$	$\log \frac{S_{f,t}^i}{S_{f,t}}$	$\log \frac{\Delta S_{f,t}^i}{\Delta S_{f,t}}$
(1)	(2)	(3)	(4)	(5)	(6)	(6)
$\text{Log}(S_{f,t-1})$	-.314*** (0.011)	-0.324*** (0.010)	-.390*** (0.040)	-0.457*** (0.025)	-0.426*** (0.012)	-0.446*** (0.010)
$\text{Log}(aw_{f,t-1}^i)$	-.012 (0.008)	-0.033*** (0.008)	-.101*** (0.012)	-0.043*** (0.007)	-0.050*** (0.005)	-0.032*** (0.004)
$\text{Log}(aw_{f,t-1}^i \times S_{f,t-1})$	0.041*** (0.003)	0.039*** (0.003)	0.023*** (0.004)	0.010*** (0.001)	0.032*** (0.002)	0.018*** (0.001)
Constant	-0.905*** (0.048)	-0.902*** (0.041)	-1.729*** (0.135)	-1.269*** (0.074)	-0.875*** (0.037)	-0.977*** (0.023)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.600	0.628	0.755	0.766	0.763	0.766
No. of observations	1066867	1066867	1247310	1247310	1190681	1190681

Notes: Observations are clustered by firm. *, ** and *** indicate 10%, 5% and 1% significance, respectively.

has one or more patents at time t . d_f and d_t are firm and year fixed effects. We only use the firms that have entered at least two technology categories. The results controlling firm fixed effects are shown in Table (A.3), which are consistent with the cross-sectional findings. A firm grows larger, its R&D efforts even out in all sectors ($\beta_1 < 0$, $\beta'_1 < 0$) and that the new sectors that a firm enters are farther away from the center of the technology space than the existing sectors ($\beta_2 < 0$, $\beta'_2 < 0$). Again the results are robust at different levels of disaggregation.

Table A.3: Choices of New Sectors and Firms Patent Stock

Variable	3-digit SIC		NClass Categories		IPC Categories	
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Log}(S_{f,t-1})$	-.036*** (0.003)	-0.072*** (0.004)	-0.069*** (0.008)	-0.112*** (0.008)	-0.124*** (0.015)	-0.251*** (0.015)
New Sector		-0.454*** (0.007)		-0.287*** (0.012)		-0.810*** (0.018)
New Sector \times $\log(S_{f,t-1})$	-0.149*** (0.002)		-0.091*** (0.003)		-0.192*** (0.004)	
Constant	-2.542*** (0.046)	-2.584*** (0.053)	-5.977*** (0.031)	-6.116*** (0.027)	-7.963*** (0.073)	-7.472*** (0.076)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.039	0.087	0.001	0.002	0.028	0.036
No. of observations	1066867	1213652	991215	991215	817762	817762

Notes: Observations are clustered by firm. *, ** and *** indicate 10%, 5% and 1% significance, respectively.

Observation 4: Let g_{ext} be the extensive growth rate attributable to new patent applications in new technological classes and g_{int} be the intensive growth rate coming from patent applications in existing classes. Define S_t^{New} as the number of patent applications in *new* classes at time t .

Therefore,

$$\begin{aligned}
g_{ext,t} &= \frac{S_{f,t}^{New}}{S_{f,t-1}} \\
g_{int,t} &= \frac{(S_{f,t} - S_{f,t}^{New}) - S_{f,t-1}}{S_{f,t-1}}
\end{aligned}$$

Table A.4 presents the results of fixed-effect regressions and Heckman two-steps procedure correcting for selection bias (as only survival firms are observed). When controlling initial firm size, the firm whose initial average technology applicability $TA_{f,t-1}$ is greater also experiences faster growth rate $g_{f,t}$. At the first stage of the Heckman procedure, higher TA_{t-1} also increases the firm's survival probability.

Table A.5 shows that both extensive and intensive firm innovation rates decrease with firm's initial number of patents but increases with the applicability of firm's initial technology mix. The Inverse Mills ratio in columns (3) and (6) is taken from the columns (5) of Table A.4. Both extensive and intensive firm innovation rate are positively related to the initial average technology applicability $TA_{f,t-1}$, but the extensive margin is more sensitive to $TA_{f,t-1}$ than the intensive margin. This shows that a central location on the technology space promotes firm growth mainly through providing prerequisite knowledge while the firm expands into new sectors.

Larger firms (firms with more employees or sales) have more patents and typically grow (innovate) at a slower rate. We investigate if firm's future innovation rate is indeed negatively correlated with the commonly used measure of firm size. We merge firm patenting data with firm data from Compustat dataset and substitute patent stock with number of employers in the above regression. The results are reported in Table A.6.

Table A.4: Firm Innovation Rate and Firm's Technology Applicability

Variable	OLS (-100%)			Heckman Correction	
	(1)	(2)	(3)	survival (4)	growth (5)
$\text{Log}(S_{f,t-1})$	-.227*** (0.002)	-0.103*** (0.001)	-0.811*** (0.009)	0.335*** (0.003)	-0.143*** (0.002)
$\text{Log}(TA_{f,t-1})$	0.073*** (0.001)	0.066*** (0.001)	0.058*** (0.001)	0.090*** (0.001)	0.128*** (0.001)
Age				-0.058*** (0.000)	
Inverse Mills ratio					1.139*** (0.020)
Constant	1.595*** (0.010)	1.176*** (0.011)	1.483*** (0.011)	1.074*** (0.014)	1.371*** (0.010)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	No	No	Yes	No	No
R^2	0.259	0.107	0.210		0.273
No. of observations	333968	424504	333968	333968	333968

Notes: Observations are clustered by sector in all columns, except that they are clustered by firm in column (2) and (5). *, ** and *** indicate 10%, 5% and 1% significance, respectively. $\text{Log}(S)$ is the log-scaled number of patents. $TA_{f,t-1}$ is the weighted authority weight of a firm's patent stock.

Table A.5: Extensive v.s. Intensive Firm Innovation Rate and Technology Applicability

Variable	Extensive Firm Growth Rate			Intensive Firm Growth Rate		
	(1)	(2)	Heckman (3)	(4)	(5)	Heckman (6)
$\text{Log}(S_{f,t-1})$	-.143*** (0.001)	-0.566*** (0.008)	-0.071*** (0.001)	-0.084*** (0.001)	-0.246*** (0.004)	-0.072*** (0.001)
$\text{Log}(TA_{f,t-1})$	0.048*** (0.001)	0.056*** (0.001)	0.095*** (0.001)	0.025*** (0.000)	0.012*** (0.000)	0.033*** (0.001)
Inverse Mills ratio			0.975*** (0.014)			0.164** (0.010)
Constant	0.960*** (0.009)	0.896*** (0.009)	0.768*** (0.009)	0.635*** (0.007)	0.587*** (0.007)	0.603*** (0.007)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	No	Yes	No	No	Yes	No
R^2	0.157	0.332	0.173	0.105	0.177	0.106
No. of observations	333968	333968	333968	333968	333968	333968

Notes: Observations are clustered by sector in all columns, except that they are clustered by firm in column (2) and (5). Years covered are 1997-2006. *, ** and *** indicate 10%, 5% and 1% significance, respectively. $\text{log}(S_{f,t-1})$ is the log-scaled number of patents. $TA_{f,t-1}$ is the (weighted) authority weight of a firm's patent portfolio.

Table A.6: Firm Innovation Rate, Firm Size and Technology Applicability (using Compustat)

Variable	OLS		Heckman Correction	
	(-100%)		survival	growth
	(1)	(2)	(3)	(4)
$\text{Log}(Emp_{f,t-1})$	-0.024*** (0.001)	-0.180*** (0.010)	0.335*** (0.003)	-0.143*** (0.002)
$\text{Log}(TA_{f,t-1})$	0.006*** (0.001)	0.005*** (0.001)	0.090*** (0.001)	0.128*** (0.001)
Age			-0.058*** (0.000)	
Inverse Mills ratio				1.139*** (0.020)
Constant	0.147*** (0.009)	0.372*** (0.019)	1.074*** (0.014)	1.371*** (0.010)
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	No	Yes	No	No
R^2	0.35	0.26		0.273
No. of observations	39949	39949	333968	333968

Notes: Observations are clustered by sector in all columns, except that they are clustered by firm in column (2). *, ** and *** indicate 10%, 5% and 1% significance, respectively. $\text{log}(Emp_{f,t-1})$ is the log-scaled number of employees. $TA_{f,t-1}$ is the (weighted) average authority weight of a firm's patent portfolio.

B Technical Appendix

B.1 An All-Sector Firm's Optimal R&D Decision

We solve the firm's R&D decision along the BGP. We adopt the guess-and-verify method to solve the all-sector firm's problem. Guess that the value of a firm is a linear combination of its accessible knowledge capital in all the sectors:

$$V(z_{f,t}) = \sum_{j=1}^K \left(v_t^j \frac{z_{f,t}^j}{n_t^j} + u_t^j \right)$$

Substituting it back to the Bellman equation, we get

$$V(z_{f,t}) = \sum_{j=1}^K \left(\pi_t^j \frac{z_{f,t}^j}{n_t^j} \right) - \sum_{i=1}^K \sum_{j=1}^K \left(R_{1f,t}^{i \leftarrow j} + R_{1f,t}^{i \leftarrow j} \right) + \quad (30)$$

$$\frac{1}{1+r} \sum_{j=1}^K (v_{t+1}^j) \frac{z_{f,t}^j + \sum_{i=1}^K \left[A^{ji} \left(z_t^j R_{f,t}^{j \leftarrow i} \right)^\alpha \left(z_{f,t}^i + \theta z_t^i \right)^{1-\alpha} \right]}{n_{t+1}^j} + u_{t+1}^j. \quad (31)$$

The first order condition with respect to $R_{f,t}^{i \leftarrow j}$ is:

$$R_{1f,t}^{i \leftarrow j} = \frac{n_t^j}{n_t^i} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} M_t^i \frac{z_{f,t}^j}{n_t^j}. \quad (32)$$

$$R_{2f,t}^{i \leftarrow j} = \frac{n_t^j}{n_t^i} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} M_t^i \frac{\theta \bar{z}_t^j}{n_t^j}. \quad (33)$$

where $\rho_t^j = \frac{1}{1+r} \frac{n_t^j}{n_{t+1}^j}$. Substituting the optimal R&D in (32) and (33) back to the Bellman equation (31), we get:

$$\begin{aligned} \sum_{j=1}^K \left(v_t^j \frac{z_{f,t}^j}{n_t^j} + u_t^j \right) &= \sum_{j=1}^K \left(\pi^j \frac{z_{f,t}^j}{n_t^j} \right) - \sum_{i=1}^K \sum_{j=1}^K \frac{n_t^j}{n_t^i} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} M_t^i \left(\frac{z_{f,t}^j + \theta \bar{z}_t^j}{n_t^j} \right) \\ &+ \frac{1}{1+r_t} \left[\sum_{j=1}^K \frac{v_{t+1}^j z_{f,t}^j}{n_{t+1}^j} + \sum_{j=1}^K \sum_{i=1}^K \frac{v_{t+1}^i}{n_{t+1}^i} \left[A^{i \leftarrow j} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} (z_{f,t}^j + \theta \bar{z}_t^j) \right] + u_{t+1}^j \right] \end{aligned}$$

Therefore,

$$\begin{aligned} u_t^j &= - \sum_{i=1}^K \frac{n_t^j}{n_t^i} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} \frac{M_t^i \theta \bar{z}_t^j}{n_t^j} + \frac{1}{1+r_t} \sum_{i=1}^K A^{i \leftarrow j} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} \theta \bar{z}_t^j \frac{v_{t+1}^i}{n_{t+1}^i} + \frac{1}{1+r_t} u_{t+1}^j \\ \frac{v_t^j}{n_t^j} &= \frac{\pi_t^j}{n_t^j} - \sum_{i=1}^K \sum_{j=1}^K \frac{n_t^j}{n_t^i} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} \frac{M_t^i}{n_t^j} + \frac{1}{1+r_t} \sum_{i=1}^K A^{i \leftarrow j} \left(\frac{A^{i \leftarrow j} \alpha \rho_t^i v_t^i}{M_t^i} \right)^{\frac{1}{1-\alpha}} \frac{v_{t+1}^i}{n_{t+1}^i} + \frac{1}{1+r_t} \frac{v_{t+1}^j}{n_{t+1}^j} \end{aligned}$$

where M_t^i is the total number of firms in sector i , $M_t^i z_t^i = n_t^i$. The transversality condition takes the form

$$\begin{aligned} \lim_{T \rightarrow \infty} \prod_{t=0}^T \left(\frac{1}{1+r_t} \right) u_T^i &= 0, \forall i \\ \lim_{T \rightarrow \infty} \prod_{t=0}^T \left(\frac{1}{1+r_t} \right) \frac{v_T^i}{n_T^i} &= 0, \forall i \end{aligned}$$

In a stationary BGP equilibrium, the measure (number) of firms in a given sector is constant $M_t^i = M^i$. The sectoral knowledge values and the application value of knowledge j to i are all constant, i.e. $v_t^i = v^i$, $u_t^i = u^i$, (to be proved later). Now we get:

$$\begin{aligned} v^j &= (1 - \rho_t^j)^{-1} \left[\pi^j + \frac{1-\alpha}{\alpha} \sum_{i=1}^K \frac{n_t^j}{n_t^i} \left(A^{i \leftarrow j} \alpha \rho_t^i v^i \right)^{\frac{1}{1-\alpha}} M^{i \frac{\alpha}{\alpha-1}} \right] \\ u^j &= (1 - \frac{1}{1+r_t})^{-1} \left[\frac{1-\alpha}{\alpha} \sum_{i=1}^K \frac{n_t^j}{n_t^i} \left(A^{i \leftarrow j} \alpha \rho_t^i v^i \right)^{\frac{1}{1-\alpha}} M^{i \frac{\alpha}{\alpha-1}} \frac{\theta \bar{z}_t^j}{n_t^j} \right]. \end{aligned}$$

To simplify the notations, define the value of sector j 's knowledge in contributing to innovations in sector i as

$$\omega^{i \leftarrow j} = \frac{1-\alpha}{\alpha} \frac{n_t^j}{n_t^i} \left(A^{i \leftarrow j} \alpha \rho_t^i v^i \right)^{\frac{1}{1-\alpha}} (M^i)^{\frac{\alpha}{\alpha-1}}$$

Substituting it back, we have

$$\begin{aligned} v^j &= (1 - \rho_t^j)^{-1} (\pi^j + \sum_{i=1}^K \omega_t^{i \leftarrow j}), \\ u^j &= (1 + \frac{1}{r}) (\sum_{i=1}^K \omega_t^{i \leftarrow j} \frac{\theta \bar{z}_t^j}{n_t^j}) \end{aligned}$$

and

$$R_{f,t}^{i \leftarrow j} \equiv R_{1f,t}^{i \leftarrow j} + R_{1f,t}^{i \leftarrow j} = \frac{\alpha}{1 - \alpha} \omega_t^{i \leftarrow j} \frac{z_{f,t}^j + \theta \bar{z}_t^j}{n_t^j}$$

To prove that $\rho_t^j, v^j, u^j, \omega_t^{ij}$ are all constants, we first need to show that the innovation rates across sectors are the same on the BGP; therefore, we need to show $\frac{n_t^j}{n_t^i} = \frac{n^j}{n^i}, \forall t$. The evolution of the number of varieties in sector i is:

$$\begin{aligned} n_{t+1}^i &= n_t^i + \int_{f \in \mathcal{F}_{i,t}} \Delta z_{f,t}^i df \\ &= n_t^i + \sum_{j=1}^K (A^{i \leftarrow j})^{\frac{1}{1-\alpha}} \left(\frac{\alpha \rho_t^i v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \left[\int_{f \in \mathcal{F}_{i,t}} (z_{f,t}^j + \theta \bar{z}_t^j) df \right] \\ &= n_t^i + \sum_{j=1}^K \left[(A^{i \leftarrow j})^{\frac{1}{1-\alpha}} \left(\frac{\alpha \beta v^i}{\gamma_t^i M^i} \right)^{\frac{\alpha}{1-\alpha}} \right] (n_t^{ji} + \theta \frac{M^i}{M^j} n_t^j) \end{aligned}$$

where $n_t^{ji} \equiv \int_{f \in \mathcal{F}_{i,t} \cap \mathcal{F}_{j,t}} z_{f,t}^j df$ represents the total number of sector j goods that are produced by firms which also produce in sector i , because not all firms in sector j is innovating and producing in sector i . The second term in the last bracket represents the total public knowledge in sector j that is used for innovation in sector i . Firms can adopt public knowledge capital from every sector when innovating, but private knowledge is limited to what sectors firms have previously entered. The innovation rate (the growth rate of varieties) in sector i is $\gamma_t^i = n_{t+1}^i / n_t^i$. Rearranging the terms, we have

$$(\gamma_t^i - 1)(\gamma_t^i)^{\frac{\alpha}{1-\alpha}} = \left(\frac{\alpha \beta v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \sum_{j=1}^K (A^{i \leftarrow j})^{\frac{1}{1-\alpha}} \left(\frac{n_t^{ji}}{n_t^i} + \theta \frac{M^i}{M^j} \frac{n_t^j}{n_t^i} \right), \quad (34)$$

The number of goods in every sector grows at the same speed, because inter-sector knowledge spillovers keep all sectors on the same track. More specifically, if one sector i had been growing more slowly than other sectors for a lengthy period, its number of goods would be extremely small relative to other sectors. (34) implies that the cross-sector knowledge spillovers would increase γ_t^i tremendously through a large ratio n_t^{ji}/n_t^i and n_t^j/n_t^i until γ_t^i is the same as the innovation rates in other sectors. This is vice versa for sectors starting with a slower growth rate. Therefore, in

the stationary BGP equilibrium, $\gamma^i = \gamma^j = \gamma$ and the distribution of the sector is stable and rank-preserving. Denote $\frac{n_t^i}{n_t^j} = \frac{n^i}{n^j}, \forall t$.

This result implies that $\rho_t^j = \beta/\gamma \equiv \rho$ and $\omega_t^{ij} \equiv \omega^{ij}$ are both constants, consistent with our original guess. Therefore, we have (13), (15), (14) and (16). Now we can verify our previous guess that the all-sector firm's value is a linear constant-coefficient combination of its knowledge in all sectors:

$$V(z_{f,t}) = \sum_{i \in \mathcal{S}_{f,t}} v^i \frac{z_{f,t}^i}{n_t^i} + u^i.$$

B.2 Sectoral Innovation Rate and Research Intensity

The number of varieties (patents) in sector i accumulates according to

$$n_{t+1}^i = n_t^i + \int \Delta z_{f,t}^i df$$

Substitute (11) into the above equation, we get

$$\begin{aligned} n_{t+1}^i &= n_t^i + \int_{f \in \mathcal{F}_{i,t}} \sum_{j=1}^K \left[(A^{i \leftarrow j})^{\frac{1}{1-\alpha}} \left(\frac{\alpha \rho_t^i v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \left(z_{f,t}^j \varepsilon_{1f,t}^{ij} + \theta \bar{z}_t^j \varepsilon_{2f,t}^{ij} \right) \right] df \\ &= n_t^i + \sum_{j=1}^K (A^{i \leftarrow j})^{\frac{1}{1-\alpha}} \left(\frac{\alpha \rho_t^i v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \int_{f \in \mathcal{F}_{i,t}} (z_{f,t}^j + \theta \bar{z}_t^j) df, \end{aligned}$$

which implies the common innovation rate is

$$\begin{aligned} \gamma &= 1 + \sum_{j=1}^K \frac{n^j}{n^i} \left[(A^{i \leftarrow j})^{\frac{1}{1-\alpha}} \left(\frac{\alpha \rho v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \right] \frac{\int_{f \in \mathcal{F}_{i,t}} (z_{f,t}^j + \theta \bar{z}_t^j) df}{n_t^i} \\ &= 1 + \sum_{j=1}^K \frac{\omega^{i \leftarrow j} \tau^{i \leftarrow j}}{(1-\alpha) \rho v^i}. \end{aligned} \tag{35}$$

where $\tau^{i \leftarrow j} \equiv \int_{f \in \mathcal{F}_{i,t}} (z_{f,t}^j + \theta \bar{z}_t^j) df / n_t^i$ stands for the fraction of knowledge in j that is utilized in innovating in i . Based on (13), we can rewrite the equation above as

$$\gamma = 1 + \frac{1-\rho}{(1-\alpha)\rho} \frac{\sum_{i=1}^K \sum_{j=1}^K \omega^{i \leftarrow j} \tau^{i \leftarrow j}}{\sum_{i=1}^K \pi^i + \sum_{i=1}^K \sum_{j=1}^K \omega^{i \leftarrow j}},$$

Substituting out $\rho = \beta/\gamma$ leads to (26) after rearranging the terms.

The sectoral research intensity is defined as the overall sectoral R&D expenditure divided by

sectoral revenue: $RI^i \equiv \frac{1}{s^i PY} \sum_{j=1}^K \int_{f \in \mathcal{F}_i \cap \mathcal{F}_j} R_f^{i \leftarrow j} df$. Substitute the optimal R&D expenditure (16) and (35) into the equation, we have

$$\begin{aligned} RI^i &= \frac{\alpha}{1-\alpha} \frac{1}{s^i PY} \sum_{j=1}^K \omega^{i \leftarrow j} \frac{\int_{f \in \mathcal{F}_i} (z_{f,t}^j + \theta^j \bar{z}_t^j) df}{n_t^j} \\ &= \frac{\alpha}{1-\alpha} \frac{1}{s^i PY} \sum_{j=1}^K \omega^{i \leftarrow j} \tau^{i \leftarrow j} \\ &= \frac{\alpha \rho (\gamma - 1) K}{PY} \frac{v^i}{v^j} \end{aligned}$$

Therefore, $RI^i/RI^j = v^i/v^j$.

B.3 The Evolution of (Normalized) Firm Size

Based on knowledge accumulation (10), knowledge production (11) and optimal R&D investment (16), firm f accumulates its knowledge in sector i according to

$$\begin{aligned} z_{f,t+1}^i &= z_{f,t}^i + \sum_{j=1}^K [A^{i \leftarrow j} (\bar{z}_t^i R_{1f,t}^{i \leftarrow j})^\alpha (z_{f,t}^j)^{1-\alpha} \varepsilon_{1f,t}^{ij} + \theta A^{i \leftarrow j} (\bar{z}_t^i R_{2f,t}^{i \leftarrow j})^\alpha (\bar{z}_t^j)^{1-\alpha} \varepsilon_{2f,t}^{ij}] \\ &= z_{f,t}^i + \sum_{j=1}^K A^{i \leftarrow j} \left(\bar{z}_t^i \frac{\alpha}{1-\alpha} \frac{\omega^{i \leftarrow j}}{n_t^j} \right)^\alpha (z_{f,t}^j \varepsilon_{1f,t}^{ij} + \theta \bar{z}_t^j \varepsilon_{2f,t}^{ij}) \end{aligned}$$

Divide both sides by n_{t+1}^i , we can write the dynamics of firm market share as

$$\begin{aligned} \frac{z_{f,t+1}^i}{n_{t+1}^i} &= \frac{n_t^i}{n_{t+1}^i} \frac{z_{f,t}^i}{n_t^i} + \frac{n_t^i}{n_{t+1}^i} \sum_j \frac{z_{f,t}^j}{n_t^j} \frac{n^j}{n^i} \left[A^{ij} \left(\frac{A^{i \leftarrow j} \alpha \rho v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \varepsilon_{1f,t}^{ij} \right] + \frac{n_t^i}{n_{t+1}^i} \sum_{j=1}^K \frac{\theta \bar{z}_t^j}{n_t^i} \left[A^{ij} \left(\frac{A^{i \leftarrow j} \alpha \rho v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \varepsilon_{2f,t}^{ij} \right] \\ &= \frac{1}{\gamma} \left[\frac{z_{f,t}^i}{n_t^i} + \sum_j \frac{z_{f,t}^j}{n_t^j} \frac{n^j}{n^i} (A^{i \leftarrow j} \left(\frac{A^{ij} \alpha \rho v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \varepsilon_{1f,t}^{ij}) \right] + \frac{1}{\gamma} \sum_{j=1}^K \frac{\theta^j}{M^j} \frac{n^j}{n^i} \left[A^{i \leftarrow j} \left(\frac{A^{i \leftarrow j} \alpha \rho v^i}{M^i} \right)^{\frac{\alpha}{1-\alpha}} \varepsilon_{2f,t}^{ij} \right] \\ &= \frac{1}{\gamma} \left[\frac{z_{f,t}^i}{n_t^i} + \sum_j \frac{z_{f,t}^j}{n_t^j} \left(\frac{\omega^{i \leftarrow j}}{(1-\alpha) \rho v^i} \varepsilon_{1f,t}^{ij} \right) \right] + \frac{1}{\gamma} \sum_{j=1}^K \frac{\theta^j}{M^j} \frac{\omega^{i \leftarrow j}}{(1-\alpha) \rho v^i} \varepsilon_{2f,t}^{ij} \end{aligned}$$

Define $\phi_{f,t}^{ij} \equiv \frac{1}{\gamma} (1_{\{i=j\}} + \xi^{ij} \varepsilon_{1f,t}^{ij})$, $\psi_{f,t}^{ij} \equiv \frac{\xi^{ij} \varepsilon_{2f,t}^{ij}}{\gamma}$, we can rewrite the above equation as in (24).