October 2013
Comments welcome

Buy, Keep or Sell:

Market for Ideas and Economic Growth

by

Ufuk Akcigit, Murat Alp Celik and Jeremy Greenwood[*] [†]

Abstract

An endogenous growth model is developed where each period firms invest in researching and developing new ideas. An idea increases a firm's productivity. By how much depends on how central the idea is to a firm's activity. Ideas can be bought and sold on a market for patents. A firm can sell an idea that is not relevant to its business or buy one if it fails to innovate. The developed model is matched up with stylized facts about the market for patents in the U.S. The analysis attempts to gauge how efficiency in the patent market affects growth.

Keywords: Growth, Ideas, Innovation, Misallocation, Patents, Patent Agents, Research and Development, Search frictions

JEL Nos: O31, O41

[*] *Affiliation*: University of Pennsylvania.

[†] Address all correspondence to Ufuk Akcigit: uakcigit@econ.upenn.edu.

1 Introduction

1.1 Motivation

New ideas are the engines of economic progress. The rise in living standards depends on the effectiveness of transforming new ideas into consumer products or production processes. Incarnating an idea into a product or a production process is by no means immediate. Someone must have a vision or application for the idea and the know-how to implement it. These are often people who work in areas related to the end-use of an idea.

For example, in 1849 Walter Hunt was granted a patent for the safety pin. In the abstract for the patent, Walter Hunt wrote "(t)he distinguishing feature of the invention consist in the construction of a pin made of a piece of wire or metal combining a spring, and a clasp or catch, in which catch the point of the said pin is forced and by its own spring securely retained"–see Figure 1 for the drawings included with his patent application. Hunt was a mechanic by trade and filed patents for various things, such as ice boats, machines for cutting nails, repeating guns. In need of cash, Hunt sold his patent to W. R. Grace and Company for about $10,000 (in today's dollars). W. R. Grace and Company massed produced the safety pin and made millions.

Walter Hunt by no means was an exception. Recently released data on U.S. secondary markets for patents indicate that a large fraction of patents are sold by firms who developed the idea to other firms. Specifically, among all the patents registered in the USPTO, 16% are traded and this number goes up to 20% among domestic patents; see the right panel of Figure 2. For economic progress, not only the possibility of exchange, but also the speed of that process is important. USPTO data shows that patents are sold among firms on average within 5.34 years (with a standard deviation of 4.38 years); the left panel of Figure 2 shows the frequency distribution over the duration for a sale. So, it takes time to sell a patent.

Firms often develop patents that are not close to their primary business activity. Think about a patent as lying within some technological class. Call this technology class j. Empirically this can be represented by the first two digits of its IPC code. Now, one can measure

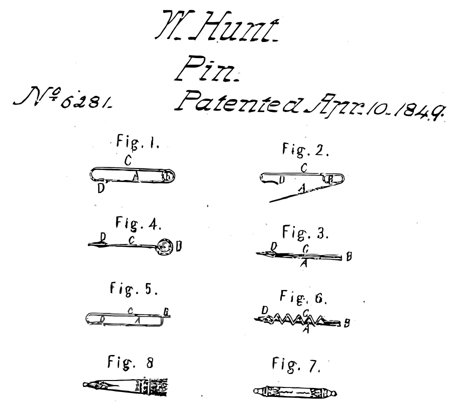

Figure 1: Patent US6281. The drawings for the safety pin in Walter Hunt's patent application.

how close two patents classes, j and k, are to each other. To do this, let $\#(X \cap Y)$ denote the number of all patents that cite patents from technology classes X and Y simultaneously. Let $\#(X \cup Y)$ denote the number of all patents which cite either technology class X or Y or both. Then the following symmetric distance metric can be constructed:

$$d(X,Y) \equiv 1 - \frac{\#(X \cap Y)}{\#(X \cup Y)}, \text{ with } 0 \leq d(X,Y) \leq 1.$$

This distance metric is intuitive. If each patent that cites X also cites Y, this metric delivers a distance of $d(X,Y) = 0$. If there is no patent that cites both classes, then the distance becomes $d(X,Y) = 1$. The distance between two technology classes increases, as the fraction of patents that cite both decreases. Given this metric between technology classes, a distance measure between a patent and a firm can now be constructed.

In order to measure how close a patent is to a firm in the technology spectrum, a metric needs to be devised. For this purpose, a firm's past patent portfolio is used to identify the firm's existing location in the technology space. In particular, the distance of a particular patent p to a firm f is computed by calculating the average distance of p to each patent in

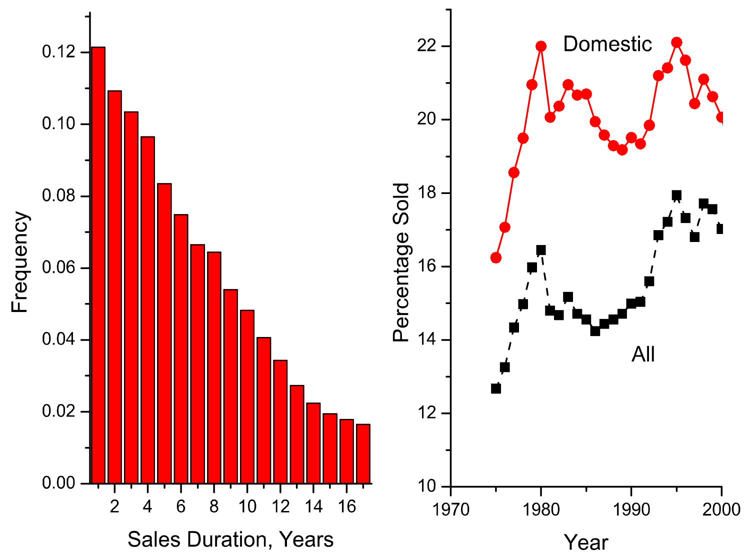

Figure 2: Duration and Sales.

firm f's patent portfolio as follows:

$$d_\iota\left(p,f\right)\equiv[\frac{1}{\|\mathcal{P}_f\|}\sum_{p'\in\mathcal{P}_f}d\left(X_p,Y_{p'}\right)^{\iota}]^{1/\iota},\ \text{with } 0<\iota\leq 1,\ \text{and where } 0\leq d_\iota(p,f)\leq 1. \tag{1}$$

In this expression, $\mathcal{P}_f$ denotes the set existing patents of firm f prior to patent p, $\|\mathcal{P}_f\|$ stands for its cardinality, and $d\left(X_p,Y_{p'}\right)$ measures the distance between the technology classes of patents p and p'. Note that $d\left(X_p,Y_{p'}\right)=0$ when the firm has other patents in the same class as p. Therefore, this metric is defined only for $\iota>0$. Finally, when $\iota=1$ the above metric returns the average distance of p to each patent in firm f's patent portfolio:

$$d_1\left(p,f\right)\equiv\frac{1}{\|\mathcal{P}_f\|}\sum_{p'\in\mathcal{P}_f}d\left(X_p,Y_{p'}\right),\ \text{with } 0\leq d_1(p,f)\leq 1.$$

The empirical distribution for this notion of distance is displayed in Figure 3, for three values of ι. As can be seen, patents have heterogeneous technological distances to the inventing firms. An analysis of the patent data in Section 3 uncovers some additional important facts about the nature of these exchanges. In particular:

1. A patent contributes more to a firm's sales and stock market values if it is closer to the firm in terms of technological distance.

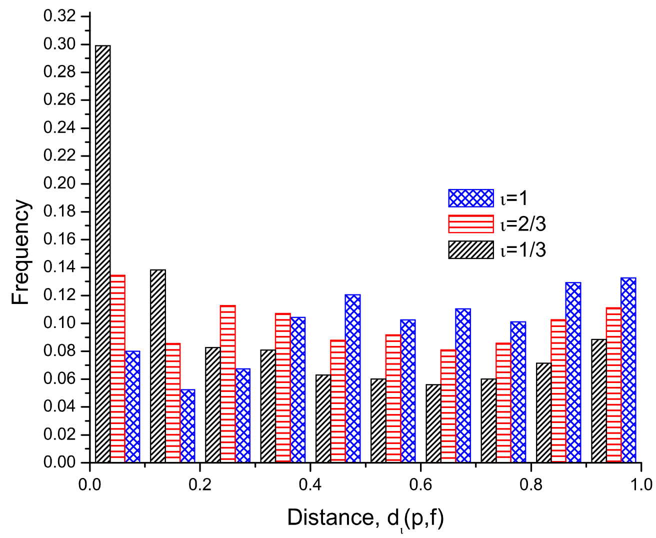

Figure 3: Empirical distance distributions. The figure plots empirical density functions for the distance, $d_\iota(p, f)$, between a patent, p, and a firm's patent portfolio, f, for three values of ι.

2. A patent is more likely to be sold the more distant it is to the inventing firm.

3. A patent is technologically closer to the buying firm than to the selling firm.

These listed facts, in conjunction with those displayed in Figures 2 and 3, raise important questions that have been left unanswered by the existing literature: How sizeable is the misallocation of ideas? How does the secondary market for ideas affect economic growth? Do frictions in the secondary market lead to more in-house R&D or do they discourage innovations overall? This paper is an attempt to answer these questions.

1.2 The Analysis

To do this, a search-theoretic model of the market for patents is built here. Each period firms invest in research and development. Sometimes this process generates an idea, other times it doesn't. Each firm operates within a particular technology class. An idea increases a firm's productivity. The extent to which it does depends on the proximity of the idea to the firm's technology class. A firm may wish to sell an idea that isn't close to its own class. It can do so using a patent agent. Analogously, the firm might want to purchase an idea through a patent agent if it fails to innovate. Due to search frictions it may take time for a patent agent to find a buyer for a patent. Also, a patent may not be the perfect match for a buyer. Due to R&D by firms there is growth in the model. A balanced growth path for the model is explicitly characterized. A unique invariant firm-size distribution exists despite the fact that the distribution for productivity across firms is continually fanning out.

The model is calibrated so that it matches certain features of the U.S. economy, such as the average rate of growth, the share of R&D in GDP, the share of patents that are sold, the average duration of a sale, the spread in sale time, etc. Clearly, a market for patents affects the incentive to do R&D. On the one hand, the fact that an idea, which is not so useful for the innovator's own production, can be sold raises the return from engaging in R&D. On the other hand, the fact that a firm can buy an idea reduces the reward from doing R&D. A goal of the analysis is to examine how a patent market affects R&D and, hence, growth.

To gauge the importance of the patent resale market for economic growth and welfare, a sequence of structured thought experiments is undertaken. In this sequence of thought experiments the efficiency of the patent market is successively increased. First, it is shown that a faster rate of contact between buyers and sellers leads to a welfare gain of 11% (measured in terms of consumption). Next, if a seller's patent could be perfectly matched with a buyer's ideal idea then a further welfare of 36% materializes. Last, if the ideas that firms produce are perfectly suited for their own production process (this corresponds to a situation where there is no mismatch between a firm and the idea that it generates) then welfare would be 50% higher than in the baseline model. So, it seems that efficiency in the resale market for patents does matter.

The market for patents is often thought of as being inefficient and illiquid. Buying and selling intellectual property is a difficult activity. Each patent is unique. It may not be readily apparent who the potential buyers and competing sellers even are, especially in situations where enterprises desire to keep their business strategies secret. Buyers and sellers may have very different valuations about the worth of a patent. Patents are often sold through intermediaries. This motivates the search-theoretic framework presented here.

Historically patent agents were often lawyers. Dealing with patent buyers and sellers they understood both sides of the market. Inventors used them to file patent applications. So, the lawyers became acquainted with the new technologies that were around. Buyers used them to vet the merits of new technologies. Hence, the lawyers were familiar with the types of patents that were likely to be marketable. This lead naturally to the lawyers acting as intermediaries in patent sales. Edward Van Winkle typifies the business. He was a patent agent at the beginning of the 20th century. Van Winkle was a mechanical engineer who acquired a law degree by correspondence course. He was well suited to provide advice on the legal and technical merits of inventions for his clients on both sides of the market. Van Winkle cultivated a network of businessmen, inventors, and other lawyers. Lamoreaux and Sokoloff (2003) detail how he brokered various types of deals with the buyers and sellers of patents. They also document for the period 1870 to 1910 an increased tendency for inventors (especially the more productive ones) to use specialized registered patent agents to handle transactions associated with their patents.

Even today the market for patents is thin, according to Gans and Stern (2010) and Hagiu and Yoffie (2011). The patent market is highly specialized. Gans and Stern (2010) and Hagiu and Yoffie (2011) discuss the failure of online intellectual property platforms to arbitrage the market. According to them, the sensitivity of intellectual property makes potential buyers and sellers reluctant to reveal information online; they prefer face-to-face dealings with the other party. Also, some buyers may perceive a lemon's problem: if the patents were truly valuable then the seller should be able to profit by developing the idea themselves or by selling it directly to interested parties.[1]

---

[1] The current analysis ignores the rise of non-practicing entities and defensive aggregators. The former buy patents and then seek licensing fees from firms under threat of litigation. The latter try to reduce litigation for firms by buying "toxic patents" and providing licences. There also firms that are hybrids between these two forms of entities. See Hagiu and Yoffie (2011) for a discussion.

1.3 Relationship to the Literature

The paper contributes to a few strands in the literature. The research builds and extends models of endogenous growth with quality improvements. [See Aghion, Akcigit and Howitt (2013), and the references therein, for a recent survey of Schumpeterian growth models]. Recently attention has been directed to how new ideas spread in an economy. Some work stresses technology diffusion via innovation and imitation [e.g., Acemoglu, Aghion and Zilibotti (2006), Jovanovic and MacDonald (1994), and Konig, Lorenz and Zilibotti (2012)]. Other work emphasizes matching and other frictions in the transfer of ideas. [See for instance, Benhabib, Perla, and Tonetti (2012), Chatterjee and Rossi-Hansberg (2012), Chiu, Meh and Wright (2011), and Lucas and Moll (2011)].

The work here emphasizes matching frictions. It differs from the above papers in a number of significant ways: First, the focus is on an economy where growth is driven by *heterogenous* ideas that are invented by firms. A firm may not be able to make the best use of the idea discovers. Second, firms can trade their ideas in a *secondary market* subject to the matching frictions. Third, while the literature has mainly been theoretical, the current research uses micro data on *patent reassignments* to motivate and discipline the analysis. Last, on the empirical side, the paper is related to interesting empirical work by Serrano (2010, 2011). It contributes to this empirical line of work by introducing some new facts on the transfer of ideas. It also illustrates how patent data can be used in conjunction with an endogenous growth model to gauge the extent of mismatch in ideas that may arise in an economy.

The focus on mismatch in ideas connects with recent work on misallocation [see for instance, Acemoglu, Akcigit, Bloom, and Kerr (2013), Guner, Ventura, and Xu (2008), Hsieh and Klenow (2009), and Restuccia and Rogerson (2008)]. Ideas are not necessarily born to their best users. The existence of a secondary market for ideas and its efficiency can have a major impact on mitigating any initial misallocation. Thus, the presence of a secondary market may contribute significantly to productivity growth. Addressing this question is the focus of the current paper.

2 Model

2.1 Environment

Consider an economy, where time flows discretely, with a continuum of firms of unit measure. The firms produce a homogeneous final good using capital and labor. Each firm belongs to some technology class j that resides on a circle with radius $1/\pi$. At each point on the technology circle there are firms of mass $1/2$. A firm enters the period with a level of productivity z. At the beginning of a period each firm develops an innovation with probability i. The innovation will be patented and belongs to some technology class k on the circle. The distance between the firm's own technology class, j, and the innovation, k, is given by $d(j,k)$. This represents the length of the shortest arc between j and k. Transform this distance measure into a measure of technological proximity, $x = 1 - d(j,k)$, defined on $[0,1]$. A high value for x indicates that the innovation is close to the firm's technology class. The value of x is drawn from the distribution function $X(x)$. The technology circle is illustrated in the left panel of Figure 4. The analysis will focus on symmetric equilibrium around this circle.

Firms produce output, o, at the end of a period according to the production process

$$o = (z')^{\zeta} k^{\kappa} l^{\lambda}, \text{ with } \zeta + \kappa + \lambda = 1, \tag{2}$$

where k and l are the amounts of capital and labor used in production and z' is its end-of-period productivity. Labor is hired at the wage rate w. There is one unit of labor available in the economy. Capital is hired at the rental rate $\widetilde{r}$.

Now, at the beginning of a period firms pick the probability of a successful innovation, i. They do this according to the convex cost function

$$C(i; \mathbf{s}) = \chi \mathbf{z}^{\zeta/(\zeta+\lambda)} i^{1+\rho}/(1+\rho). \tag{3}$$

where $\mathbf{z}$ is the mean of the productivity distribution in the economy at the beginning of the period. Cost rises in lock-step fashion with average productivity, $\mathbf{z}$, in the economy, which will be proportional to wages. (It will be established later that wages will grow at the same rate as $\mathbf{z}^{\zeta/(\zeta+\lambda)}$.) Aggregate productivity will be a function of the aggregate state of the world represented by $\mathbf{s}$. A precise definition for the aggregate state of the world will be provided later. A firm that successfully innovates can either keep or sell its idea to a patent agent. A firm that does not innovate can try to buy a patent from an agent. A patent on the market survives over time with probability σ. After it expires it cannot be used, an assumption made for technical convenience.

Incorporating a patent of closeness x increases a firm's end-of-period productivity, z', according to the law of motion

$$z' = L(z, x; \mathbf{s}) = z + \gamma x \mathbf{z}, \tag{4}$$

where z is the initial productivity level and x is the technological proximity of the patent to the firm. Two things to note about this law of motion. First, the closer is an innovation to a firm's own technology class, as represented by x, the bigger will be the increase in productivity, $z' - z$. Second, the higher is the economy-wide baseline level of productivity, $\mathbf{z}$, the more valuable a patent will be for increasing productivity. In equilibrium, $\mathbf{z}$ will be a function of the aggregate state of the world as denoted by $\mathbf{s}$.

A firm that fails to innovate can try to buy a patent from a patent agent. Likewise, a firm that draws an innovation may sell the associated patent to a patent agent at the fixed price q. This price is determined on a competitive market. Once a patent is sold to an agent the seller cannot use it in the future. A patent agent can only handle one patent at a time. The introduction of patent agents simplifies the analysis. Without this construct the analysis would have to keep track of the portfolio of patents that each firm has for sale. This technical construct is imposed without apology, as in the real world many patents are sold through agents, as was discussed. Additionally, buying or sell a patent can be thought of being equivalent to buying or selling an exclusive licensing arrangement for the idea. Licensing arrangements are inferior to patents in some respects. The purchaser of the licensing arrangement cannot sue parties that infringe on the underlying patent because the former does not own the patent. A non-exclusive licensing arrangement may leave the purchaser facing competition from other license holders. This may be an important limitation for firms selling products in national or international markets, where they may certainly be competitors.

Let n_a and n_b represent the numbers of agents and buyers in a market. The total number of matches in the market is given by the matching function

$$M(n_a, n_b) = \eta(n_a)^{\mu}(n_b)^{1-\mu}.$$

The matches are completely random. Thus, the odds that an agent will find a buyer are given by

$$m_a(\frac{n_a}{n_b}) = \frac{M(n_a, n_b)}{n_a} = \eta(\frac{n_b}{n_a})^{1-\mu},$$

and similarly that a buyer will find an agent by

$$m_b(\frac{n_a}{n_b}) = \frac{M(n_a, n_b)}{n_b} = \eta(\frac{n_a}{n_b})^{\mu}.$$

Since agents and buyers are matched randomly, the proximity between the buyer's technology class and the class of the patent being sold is a random variable. A buyer will incorporate a patent that he purchases into his production process in accordance with the above law of motion for z. The price of the patent is determined by Nash bargaining between the agent and buyer.[2] Represent this price by $p = P(z, x; \mathbf{s})$. The negotiated price will depend on the proximity of the patent, x, and the state of the buyer's technology, z. The bargaining power of the agent is given by ω. In contrast, the price at which a firm sells its patent to an agent is fixed at q, because the agent doesn't know who he will sell the patent to in the future. The timing of events is portrayed in the right panel of Figure 4.

2.2 The Representative Consumer/Worker

In the background of analysis is a representative consumer/worker. This individual supplies one unit of labor inelastically. The person owns all of the firms in the economy. He also rents

[2] In the current analysis, buyers know the quality of the match of the patent that they are purchasing. Chatterjee and Rossi-Hansberg (2012) examine the situation where the quality of an idea is private information. In such a setting there is a lemons problem in the market for ideas. It may pay for developers to start up new companies to implement good ideas because they cannot be sold at a favorable price. Such considerations are absent here.

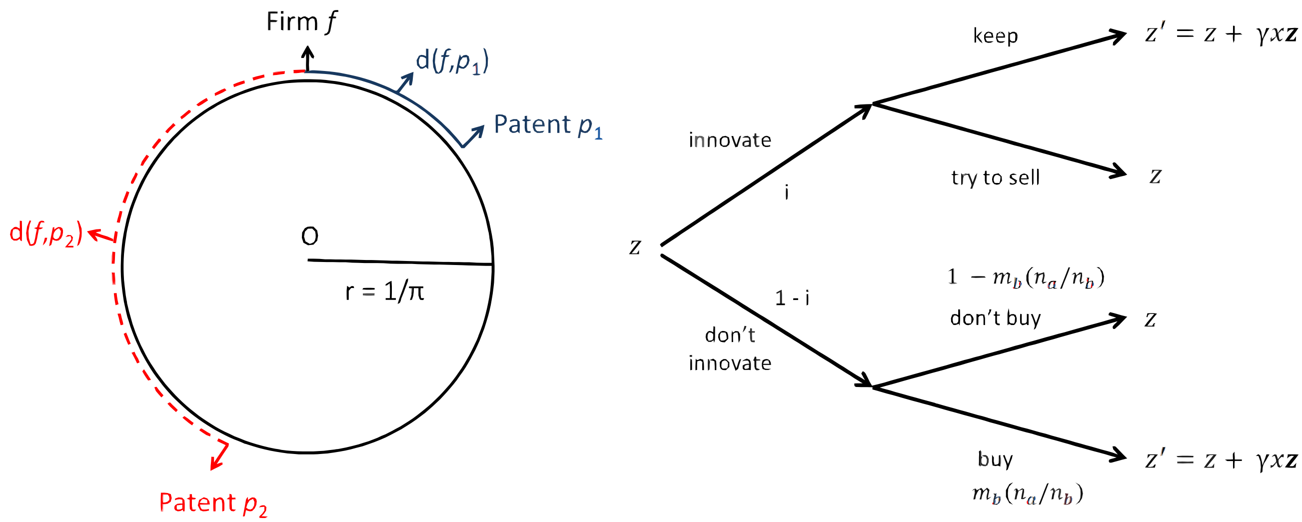

Figure 4: Technology circle (left panel) and the timing of events (right panel).

out the capital used by firms. Thus, he will earn income from wages, profits and rentals. Capital depreciates at the rate δ. The real return earned by renting capital is $1/r$. (I.e., r is the reciprocal of the gross interest rate. It will play the role of the discount factor in the Bellman equations formulated below.) The individual is assumed to have a momentary utility function of the form $U(c) = c^{1-\varepsilon}/(1-\varepsilon)$, where c is his consumption in the current period and ε is the coefficient of relative risk aversion. He discounts the future at rate β. Last, the representative consumer/worker's goal in life is to maximize his discounted lifetime utility. Since this problem is entirely standard it is not presented.

2.3 Firms

A producer hires labor, l, at the wage rate w, and capital, k, at the rental rate, $\widetilde{r} = 1/r - (1-\delta)$, to maximize profits. Thus, its decision problem is

$$\Pi(z'; \mathbf{s}) = \max_{k,l}[(z')^{\zeta} k^{\kappa} l^{\lambda} - \widetilde{r}k - wl],$$

where $\Pi(z';\mathbf{s})$ is the profit function associated with the maximization problem. The first-order conditions to this maximization problem imply that

$$k = \kappa \frac{o}{\widetilde{r}}, \tag{5}$$

and

$$l = \lambda \frac{o}{w}. \tag{6}$$

Using (2), (5) and (6), it follows that profits are given by

$$\Pi(z';\mathbf{s}) = (1-\kappa-\lambda)o = z'(1-\kappa-\lambda)[(\frac{\kappa}{\widetilde{r}})^{\kappa}(\frac{\lambda}{w})^{\lambda}]^{1/\zeta}. \tag{7}$$

Again, in equilibrium the rental and wage rates will be functions of the aggregate state of the world.

Let $V(z;\mathbf{s})$ represent the expected present-value of a firm that has a technology level of z and is about to learn whether or not it has successfully innovated. Due to the focus on symmetric equilibrium there is no need ever record the firm's location on the technology circle. Now, suppose that the firm does not innovate. Then, it will try to buy a patent. With probability $m_b(x)$ it will meet an agent selling a patent in technology class x. The patent sells at the price $p = P(z,x;\mathbf{s})$, which is a function of the buyer's type, z, as well proximity of the patent to the firm's technology class, x. The firm will only buy the patent if it yields a higher payoff than what it will obtain if it doesn't buy it. The determination of the patent price is discussed below. Denote the distribution, over proximity, for the patent agents by $D(x)$. The expected discounted present value of the buyer, $B(z;\mathbf{s})$, is

$$\begin{aligned} B(z;\mathbf{s}) \quad = \quad & m_b(\frac{n_a}{n_b}) \int \{I_a(z,x;\mathbf{s})[\Pi(L(z,x;\mathbf{s});\mathbf{s}) - P(z,x;\mathbf{s}) + rV(L(z,x;\mathbf{s});\mathbf{s}')] \\ & +[1-I_a(z,x;\mathbf{s})][\Pi(z;\mathbf{s}) + rV(z;\mathbf{s}')]\}dD(x) \\ & +[1-m_b(\frac{n_a}{n_b})][\Pi(z;\mathbf{s}) + rV(z;\mathbf{s}')], \end{aligned} \tag{8}$$

where

$$I_a(z,x;\mathbf{s}) = \begin{cases} 1, & \text{if the buyer purchases a patent,} \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The indicator function $I_a(z,x;\mathbf{s})$ is defined below. In the model developed here, there is no uncertainty at the aggregate level. Thus, next-period's aggregate state of the world, $\mathbf{s}'$, is just some determinisitic function of this period's aggregate state of the world, $\mathbf{s}$. The law of motion governing this dependence is omitted for convenience.

Turn now to the situation where the firm successfully innovates. If it decides to keep the patent, then the firm will have value $K(L(z,x;\mathbf{s});\mathbf{s})$, as given by

$$K(L(z,x;\mathbf{s});\mathbf{s}) = \Pi(L(z,x;\mathbf{s});\mathbf{s}) + rV(L(z,x;\mathbf{s});\mathbf{s}'). \tag{10}$$

Alternatively, it can sell the patent to an agent. The value of a seller, $S(z;\mathbf{s})$, is

$$S(z;\mathbf{s}) = \Pi(z;\mathbf{s}) + q + rV(z;\mathbf{s}'). \tag{11}$$

Once the seller puts a patent up for sale at the beginning of the period it may expire with probability $1-\sigma$. A firm that innovates will either keep or sell its patent depending on which option yields the highest value. Given this, it is easy to see that

$$I_k(z,x;\mathbf{s}) = \begin{cases} 1, & \text{if } K(L(z,x;\mathbf{s});\mathbf{s}) > S(z;\mathbf{s}), \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

2.3.1 The Decision to Innovate

The firm's decision to innovate is now cast. With probability i the firm innovates and with probability $1-i$ it doesn't. The firm chooses the probability of innovation subject to the convex cost function $C(i;\mathbf{s})$. Hence, write the innovation decision as

$$\begin{aligned} V(z;\mathbf{s}) &= \max_i\{(1-i)B(z;\mathbf{s}) + i\int\{I_k(z,x;\mathbf{s})K(L(z,x;\mathbf{s});\mathbf{s}) + [1-I_k(z,x;\mathbf{s})]S(z;\mathbf{s})\}dX(x) \\ &\quad -C(i;\mathbf{s})\}. \end{aligned} \tag{13}$$

The first-order condition associated with this problem is

$$\begin{aligned} &\int\{I_k(z,x;\mathbf{s})K(L(z,x;\mathbf{s});\mathbf{s}) + [1-I_k(z,x;\mathbf{s})]S(z;\mathbf{s})\}dX(x) - B(z;\mathbf{s}) \\ = \; & C_1(i;\mathbf{s}), \end{aligned}$$

so that

$$i = R(z;\mathbf{s}) = C_1^{-1}\left(\int \{I_k(z,x;\mathbf{s})K(L(z,x;\mathbf{s});\mathbf{s}) + [1 - I_k(z,x;\mathbf{s})]S(z;\mathbf{s})\}dX(x) - B(z;\mathbf{s});\mathbf{s}\right). \tag{14}$$

2.4 Patent Agents

Turn now to the problem of a patent agent. It buys a patent at the competitively determined price q. With probability $m_a(n_a/n_b)$ it will meet a potential buyer on the market and with probability $1 - m_a(n_a/n_b)$ it won't. Denote the distribution of buyers by $G(z)$. The value for an agent, A, with a patent is thus given by

$$\begin{aligned} A(\mathbf{s}) &= m_a(\frac{n_a}{n_b})\int\int I_a(z,x;\mathbf{s})P(z,x;\mathbf{s})dG(z)dD(x) \\ &+ m_a(\frac{n_a}{n_b})r\sigma A(\mathbf{s}')\int\int [1 - I_a(z,x;\mathbf{s})]dG(z)dD(x) \\ &+ [1 - m_a(\frac{n_a}{n_b})]r\sigma A(\mathbf{s}'), \end{aligned} \tag{15}$$

where $I_a(z,x;\mathbf{s})$ is specified by (9) and is defined formally shortly below. The price of a patent is determined via Nash bargaining. Specifically, p is determined in accordance with

$$\max_p\{[\Pi(L(z,x;\mathbf{s});\mathbf{s}) - p + rV(L(z,x;\mathbf{s});\mathbf{s}') - \Pi(z;\mathbf{s}) - rV(z;\mathbf{s}')]^{1-\omega} \times [p - r\sigma A(\mathbf{s}')]^{\omega}\}.$$

The first term in brackets gives the buyer's surplus. This gives difference between the value of the firm when it secures a patent and the value when it does not. The second term details the seller's surplus. In standard fashion,

$$\begin{aligned} p &= P(z,x;\mathbf{s}) = \omega[\Pi(L(z,x;\mathbf{s});\mathbf{s}) + rV(L(z,x;\mathbf{s});\mathbf{s}') \\ &-\Pi(z;\mathbf{s}) - rV(z;\mathbf{s}')] + (1-\omega)r\sigma A(\mathbf{s}'), \end{aligned} \tag{16}$$

whenever both the buyer's and seller's surpluses are positive. The price lies between $r\sigma A(\mathbf{s}')$ and $\Pi(L(z,x;\mathbf{s});\mathbf{s}) + rV(L(z,x;\mathbf{s});\mathbf{s}') - \Pi(z;\mathbf{s}) - rV(z;\mathbf{s}')$; if the former is above the latter then no solution exists.

Now, define $I_a(z, x; \mathbf{s})$ in the following manner:

$$I_a(z,x;\mathbf{s}) = \begin{cases} 1, & r\sigma A(\mathbf{s}') \leq p \leq \Pi(L(z,x;\mathbf{s});\mathbf{s}) + rV(L(z,x;\mathbf{s});\mathbf{s}') - \Pi(z;\mathbf{s}) - rV(z;\mathbf{s}') \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

2.5 Symmetric Equilibrium Along a Balanced Growth Path

The focus of the analysis is solely on symmetric equilibrium along a balanced growth path. In equilibrium the demand for labor must equal the supply of labor. Recall that there is one unit of labor in the economy. Let $Z'(z')$ represent the end-of-period distribution of z' across firms. Now, using (2), (5) and (6) it is easy to deduce that

$$l = (\frac{\kappa}{\widetilde{r}})^{\kappa/\zeta}(\frac{\lambda}{w})^{(\zeta+\lambda)/\zeta} z'. \tag{18}$$

Equilibrium in the labor market then implies that

$$\int (\frac{\kappa}{\widetilde{r}})^{\kappa/\zeta}(\frac{\lambda}{w})^{(\zeta+\lambda)/\zeta} z' dZ'(z') = 1,$$

so that the aggregate wage rate, w, is given by

$$w = \lambda(\frac{\kappa}{\widetilde{r}})^{\kappa/(\zeta+\lambda)}[\int z' dZ'(z')]^{\zeta/(\zeta+\lambda)}. \tag{19}$$

The wage rate, w, depends on the mean of the end-of-period productivity distribution across firms, $\int z' dZ'(z')$.

Next, suppose that there is free entry by agents into the market to buy patents from firms. This dictates that the price q will be determined by

$$q = A(\mathbf{s}). \tag{20}$$

To complete the description of an equilibrium, the evolution of the distributions for D, G and Z must be described. First, the distribution for D must be uniform in a symmetric equilibrium. Recall that a firm's location in the technology space is represented by a point on the circle. Think of a buyer located at the top of the circle. Suppose that a set of firms on some tiny arc jk to the left of top are selling patents of mass λ that are of distance between 0 and ε away from the top. Now take any other arc lm of equal length even further to the left of top. In a symmetric equilibrium there will be, for all practical purposes, an identical set of firms selling patents of mass λ that are of distance between $d(j,l)$ and $d(j,l)+\varepsilon$ away from the top.

Second, the distribution function Z evolves over time (from Z to Z') according to

$$Z' = \mathcal{T}_Z Z, \tag{21}$$

where the transition operator $\mathcal{T}_Z$ is specified later. Third, the distribution for patent buyers is described simply by

$$G(z) = \frac{\int^z [1 - R(y;\mathbf{s})]dZ(y)}{\int^\infty [1 - R(y;\mathbf{s})]dZ(y)}. \tag{22}$$

Any firm that fails to innovate will enter the market for patents.

Last, it is obvious that the aggregate level of productivity, $\mathbf{z}$, its gross rate of growth, $\mathbf{g}$, and the aggregate level of innovation, $\mathbf{i}$, are given by

$$\mathbf{z} = \int z dZ(z),\ \mathbf{g} = \frac{\int z' dZ'(z')}{\int z dZ(z)},\ \text{and } \mathbf{i} = \int R(z;\mathbf{s})dZ(z). \tag{23}$$

By now it should be clear, that the aggregate state of the economy is specified by $\mathbf{s} = Z$. From this it is easy to calculate wages, w, using (19) and (21).

To summarize:

Definition 1 *An equilibrium is described by a set of allocation rules, $I_k(z,x;\mathbf{s})$ and $I_a(z,x;\mathbf{s})$, value functions for firms, $B(z;\mathbf{s})$, $K(L(z,x;\mathbf{s});\mathbf{s})$, $S(z;\mathbf{s})$, and $V(z;\mathbf{s})$, a rate of innovation for firms, $R(z;\mathbf{s})$, a value function for the patent agent, $A(\mathbf{s})$, a set of selling and buying prices, q and $P(z,x;\mathbf{s})$, and distribution functions for sellers, $D(x)$, buyers, $G(z)$, and firm productivities, $Z(z)$, such that:*

1. *The indicator function $I_k(z,x;\mathbf{s})$ specifies, in line with (12), whether or not an innovator will keep his patent, given the value functions $K(L(z,x;\mathbf{s});\mathbf{s})$ and $S(z;\mathbf{s})$.*

2. *The indicator function $I_a(z,x;\mathbf{s})$ describes, as determined by (17), whether a sale between a buyer and a patent agent will occur, given the value functions $V(z;\mathbf{s})$ and $A(\mathbf{s})$.*

3. *The value functions for firms, $B(z;\mathbf{s})$, $K(L(z,x;\mathbf{s});\mathbf{s})$, $S(z;\mathbf{s})$, and $V(z;\mathbf{s})$, are given by (8), (10), (11) and (13), given prices, q and $P(z,x;\mathbf{s})$, and the distribution functions for sellers, $D(x)$, and firm productivities, $Z(z)$.*

4. *The rate of innovation for a firm, $i = R(z;\mathbf{s})$, is specified by (14), given the distribution functions for sellers, $D(x)$, the value functions $B(z;\mathbf{s})$, $K(L(z,x;\mathbf{s});\mathbf{s})$ and $S(z;\mathbf{s})$, and the decision rule for selling $I_k(z,x;\mathbf{s})$.*

5. *The value function for patent agents, $A(\mathbf{s})$, is given by (15), (10), (11) and (13), given the selling price for a patent, $P(z,x;\mathbf{s})$, and the distribution of buyers, $G(z)$.*

6. *The prices for selling and buying patents, q and $P(z,x;\mathbf{s})$, are determined in line with (20) and (16), given the value functions $V(z;\mathbf{s})$ and $A(\mathbf{s})$.*

7. *The distribution function $Z(z)$ evolves over time according to (21), given the functions $I_k(z,x;\mathbf{s})$ and $I_a(z,x;\mathbf{s})$. The mean level, $\mathbf{z}$, and growth, $\mathbf{g}$, of firm productivity and the aggregate rate of innovation, $\mathbf{i}$, are specified by (23). The distribution function for buyers, $G(z)$, is given by equation (22), given $\mathbf{i}$ and $Z(z)$. The distribution function, $D(x)$, for patent agents is uniform.*

The analysis is restricted to studying balanced growth paths. The solution to the above economy along a balanced growth path will now be characterized. Suppose that mean level of productivity for firms, $\mathbf{z}$, grows at the constant rate $\mathbf{g}$. Specify the variables z and $\mathbf{z}$ in transformed form so that $\widetilde{\mathbf{z}} = \mathbf{z}^{\zeta/(\zeta+\lambda)}$ and $\widetilde{z} = z/\mathbf{z}^{\lambda/(\zeta+\lambda)}$. Thus, $\widetilde{\mathbf{z}}$ grows at rate $\mathbf{g}^{\zeta/(\zeta+\lambda)}$ and, on average, so will $\widetilde{z}$. It turns out that $\widetilde{\mathbf{z}}$ (or equivalently $\mathbf{z}$) is sufficient to characterize the aggregate state of the economy along a balanced growth path.

Proposition 2 *(Balanced Growth) There exists a symmetric balanced growth path of the following form:*

1. *The interest factor and rental rate on capital are given by*

$$r = \beta/\mathbf{g}^{\varepsilon\zeta/(\zeta+\lambda)}, \tag{24}$$

and

$$\widetilde{r} = \mathbf{g}^{\varepsilon\zeta/(\zeta+\lambda)}/\beta - 1 + \delta. \tag{25}$$

Here $\mathbf{g}^{\zeta/(\zeta+\lambda)}$ is the common rate of growth in consumption, capital, output and wages.

2. *The value functions for buying, keeping and selling firms have linear forms in the state variables $\widetilde{z}$ and $\widetilde{\mathbf{z}}$. Specifically, $B(z;\mathbf{s}) = \mathfrak{b}_1\widetilde{z} + \mathfrak{b}_2\widetilde{\mathbf{z}}$, $K(L(z,x;\mathbf{s});\mathbf{s}) = \mathfrak{k}_1\widetilde{z} + \mathfrak{k}_2(x)\widetilde{\mathbf{z}}$, and $S(z;\mathbf{s}) = \mathfrak{s}_1\widetilde{z} + \mathfrak{s}_2\widetilde{\mathbf{z}}$.*

3. *The indicator function for an innovator specifies a threshold rule such that* $I_k(z, x; \mathbf{s}) = 1$,*whenever* $x > x_k$, *and is zero otherwise.*

4. *The indicator function for a sale between a buyer and seller specifies a threshold rule such that* $I_a(z, x; \mathbf{s}) = 1$, *whenever* $x > x_a$, *and is zero otherwise.*

5. *The value function for a patent agent has the linear form* $A(\mathbf{z}) = \mathfrak{a}\widetilde{\mathbf{z}}$.

6. *The beginning-of-period value function for a firm has the linear form* $V(z; \mathbf{s}) = \mathfrak{v}_1\widetilde{z} + \mathfrak{v}_2\widetilde{\mathbf{z}}$. *The constant rate of innovation for a firm is*

$$i = \mathbf{i} = \{\frac{1}{\chi}[X(x_k)\mathfrak{s}_2 + \int_{x_k} \mathfrak{k}_2(x)dX(x) - \mathfrak{b}_2]\}^{1/\rho}. \tag{26}$$

7. *The constant rate of growth for aggregate productivity is implicitly given by*

$$\mathbf{g} - \mathbf{1} = \gamma[\mathbf{i}\int_{x_k} xdX(x) + (1 - \mathbf{i})m_b(\frac{n_a}{n_b})\int_{x_a} xdx]. \tag{27}$$

8. *The prices for selling and buying patents are*

$$q = \mathfrak{a}\widetilde{\mathbf{z}},$$

and

$$P(z, x; \mathbf{s}) = \left[(1 - \omega)\sigma r\mathbf{g}^{\zeta/(\zeta+\lambda)}\mathfrak{a} + (\omega\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\gamma x\right]\widetilde{\mathbf{z}},$$

where π *is a constant.*

9. *The matching probabilities for sellers and buyers of patents are constant and implicitly defined by*

$$m_a(\frac{n_a}{n_b}) = \eta\{\frac{\{1 - \sigma[1 - m_a(\frac{n_a}{n_b})(1 - x_a)]\}(1 - \mathbf{i})}{\sigma\mathbf{i}X(x_k)}\}^{1-\mu}, \tag{28}$$

and

$$m_b(\frac{n_a}{n_b}) = \eta\{\frac{\sigma\mathbf{i}X(x_k)}{\{1 - \sigma[1 - m_a(\frac{n_a}{n_b})(1 - x_a)]\}(1 - \mathbf{i})}\}^{\mu}. \tag{29}$$

10. *The constants* $\mathfrak{a}$, $\mathfrak{b}_1$, $\mathfrak{b}_2$, $\mathfrak{k}_1$, $\mathfrak{k}_2$, π,$\mathfrak{s}_1$, $\mathfrak{s}_2$, $\mathfrak{v}_1$, $\mathfrak{v}_2$, x_a *and* x_b *are determined by a nonlinear equation system, in conjunction with equations (24), (26), (27), (28) and (29) that determine the variables* $\mathbf{g}, \mathbf{i}$,$r, m_a(n_a/n_b)$, *and* $m_b(n_a/n_b)$, *that does not involve either* $\widetilde{z}$ *or* $\widetilde{\mathbf{z}}$.

Proof. See Appendix 8.1. ■

2.5.1 Discussion

Along a balanced growth path, wages grow at the constant gross rate $\mathbf{g}^{\zeta/(\zeta+\lambda)}$, a fact evident from equation (19). So will aggregate output and profits, as can be seen from (7). The gross interest rate, $1/r$, will remain constant along balanced growth. Point (2) implies that on average the values of the firm at the buying, selling, and keeping stages also grow at the rate of growth of output. So, the relative values of a firm at these stages remain constant along a balanced growth path. Thus, it is not surprising then that the decisions to buy, sell or keep patents in terms of proximity, x, do not change over time. Hence, the function $I_k(z, x; \mathbf{s})$ does not depend on $\mathbf{z}$. It may seem surprising that the decision doesn't depend on z, either. This transpires because a firm's profits are linear in z, as (7) shows. It turns out that $\mathfrak{k}_1 = \mathfrak{s}_1$, which implies that only x is relevant (when comparing $\mathfrak{k}_1\widetilde{z} + \mathfrak{k}_2(x)\widetilde{\mathbf{z}}$ with $\mathfrak{s}_1\widetilde{z} + \mathfrak{s}_2\widetilde{\mathbf{z}}$). Likewise, the value of a patent agent also increases at rate $\mathbf{g}^{\zeta/(\zeta+\lambda)}$–point 3. Hence, equation (20) dictates that the price, q, at which a firm can sell a patent must also grow at this rate. Additionally, it is easy to see from (16) that the price at which the agent sells a patent to firms, p, will appreciate at this rate too. Note that this price does not depend on z, because given the linear form of the value function, V, only x will be relevant (when comparing $\mathfrak{v}_1 z'$ with $\mathfrak{v}_1 z$). It's easy to deduce from equation (14) that the rate of innovation, i, will be constant over time if B, K, and S grow at the same rate as aggregate productivity. Since the decisions to buy and sell patents only depend on x, it is not surprising that the number of buyers and sellers on the patent market are fixed along a balanced growth path. Last, the evolution of shape of the distribution function Z over time does not matter for the analysis. Its mean grows at rate $\mathbf{g}$, independently of any transformation in shape.

2.5.2 Stationary Firm-Size Distribution

The model will display a stationary firm-size distribution along a balanced growth, despite the fact the distribution for Z is shifting over time and may be changing shape. To see this, substitute (19) into (18), while making use of the definition in (23), to get

$$l = \frac{z}{\mathbf{z}}.$$

Thus, the amount of labor that a firm hires is proportional to its own productivity, z, relative to the mean level of productivity in the economy, $\mathbf{z}$. So, specifying the firm-size distribution amounts to characterizing the distribution for $z/\mathbf{z}$.

To do this, focus on a firm's draw for x. This is an independently and identically distributed random variable. To see this, note that in the current setting a firm will innovate with probability i. If it innovates then it will draw x from the distribution $X[0,1]$. Conditional on innovating, it will sell the technology with probability x_k and will keep it with probability $1 - x_k$. If it fails to innovate then it will go onto the market for patents. Conditional on failing to innovate, it finds a buyer with probability m_b. When it finds a buyer then it will draw from $\mathcal{U}[0,1]$. A purchase then occurs with probability $1 - x_a$. Hence,

$$x \text{ is } \begin{cases} = 0, & \text{with } \Pr\{(1-i)[(1-m_b) + m_b x_a] + i x_k\}, \\ \sim X[x_k, 1], & \text{with } \Pr[i(1-x_k)], \\ \sim \mathcal{U}[x_a, 1], & \text{with } \Pr[(1-i)m_b(1-x_a)]. \end{cases} \tag{30}$$

Note that $0 < E[x] < 1$.

Turn to the firm's law of motion for z or (4). Divide z through by $\mathbf{z}$ to get

$$\frac{z'}{\mathbf{z}'} = \frac{\mathbf{z}}{\mathbf{z}'}\frac{z}{\mathbf{z}} + \gamma\frac{\mathbf{z}}{\mathbf{z}'}x,$$

or

$$\widehat{z}' = \underbrace{\frac{1}{\mathbf{g}}}_{<1}\widehat{z} + \gamma\frac{1}{\mathbf{g}}x, \tag{31}$$

where $\widehat{z} \equiv z/\mathbf{z}$. This is a stationary autoregressive process with a non-Gaussian error term. Here, treat $\mathbf{g}$ as a constant, The gross growth rate, $\mathbf{g}$, can be taken as a constant because it can be solved for independently of the form for the stationary distribution. Proposition 2 establishes this. In a similar vein, m_b, x_a, and x_k are known constants that are independent of the form of the stationary distribution; again, this is a consequence of Proposition 2. Note that the process for $\widehat{z}'$ will be trapped within the compact set $[0, \overline{z}]$, where $\overline{z} \equiv \gamma/[(\mathbf{g} - \mathbf{1})]$, provided that it starts off within this interval.

Proposition 3 *(Existence of a Unique Stationary Firm-Size Distribution). The stochastic process (31) converges weakly to a unique invariant distribution.*

Proof. See Appendix 8.2. ■

Denote the stationary distribution for $\widehat{z}'$ by $\widehat{Z}$. Why does the stationary firm-size distribution have a finite upper bound, $\overline{z}$? The answer is that it is difficult for a firm's productivity, z, to grow faster than aggregate productivity, $\mathbf{z}$. Growth in aggregate productivity pulls all firms along as is evident in (4). When a firm's growth in productivity pulls ahead of aggregate growth it loses this slipstream effect, so to speak. Since z' increases in an arithmetic fashion with x growth must decay when $\mathbf{z}$ is held fixed. The distribution for productivity across firms, or Z, is not stationary. Consider a point along a balanced growth path where $\mathbf{z} = 1$. The z's will be distributed on $[0, \overline{z}]$ according to $\widehat{Z}$, where $\overline{z} = \gamma/[(\mathbf{g} - \mathbf{1})]$. Next period $\mathbf{z}$ will have grown to $\mathbf{z}' = \mathbf{g}\mathbf{z}$. Now, z will be distributed on the $[0, \mathbf{g}\overline{z}]$ according to $Z' = \widehat{Z}(z/\mathbf{g})$. Note the distribution for the z's is getting stretched rightward over time; i.e., the cumulative distribution function is being defined over an ever increasing domain. In general, if in the current period $Z : [0, z] \to [0, 1]$ then for next period $Z' : [0, \mathbf{g}z] \to [0, 1]$, where $Z'(z) = Z(z/g)$. This implicitly defines the transition operator $\mathcal{T}_Z$ in (21).

3 Empirical Analysis

The empirical analysis is undertaken here. The next section provides some details on data sources and variable constructions. For further details, please see Section 9.

3.1 Data Sources

NBER-USPTO Utility Patents Grant Data (PDP). The core of the empirical analysis draws from the NBER-USPTO Patent Grant Database (PDP). Patents are exclusionary rights, granted by national patent offices, to protect a patent holder for a certain amount of time, conditional on sharing the details of the invention. PDP data contains detailed information on 3,210,361 utility patents granted by the US Patent and Trademark Office between the years 1976 and 2006. A patent has to cite another patent when the former has a content related to the latter. When patent A cites patent B, this particular citation becomes both a *backward* citation made by A to B and a *forward* citation received by B from A. Moreover, the PDP contains an International Patent Classification (IPC) code for each patent that helps identify their relevant technologies.[3] Extensive use of the forward and backward citations are made, as well as the IPC codes assigned to each patent to determine their location in technology space, the distances between technology classes and also to proxy for patent qualities. The exact methodology followed to construct these measures will be detailed below.

Patent Reassignment Data (PRD). The second source of data comes from the recently-released USPTO patent assignment files retrieved from Google Patents Beta. This dataset provides detailed information on the changes of ownership of the patents for the years 1980 to 2011. The records include 966,427 patent reassignments not only due to *sales*, but also due to *mergers, license grants, splits, mortgages, collaterals, conversions, internal transfers*, etc. Reassignment records are classified according to a search algorithm that looks for keywords, such as "assignment", "purchase", "sale", and "merger", and assigns them to their respective categories. Through this process, 99% of the transaction records are classified into their respective groups.

Compustat North American Fundamentals (Annual). In order to assess the impact of patents and their technological distance on firm moments such as sales growth and stock market valuations, the PDP patent data is linked to Compustat firms. The focus is on the balance sheets of Compustat firms between the years 1974-2006, retrieved from Wharton Research Data Services. The Compustat database and the NBER PDP database are connected using the matching procedure provided in the PDP data.

The empirical analysis requires the construction of a notion of distance in the technology space. For that purpose, the citation patterns across IPC technology fields are utilized. PDP contains the full list of citations with the identity of citing and cited patents. Since the data also contains the IPC code of each patent, the percentage of outgoing citations from one

[3] USPTO originally assigns each patent to a particular US Patent Classification (USPC) which is a system used by the USPTO to organize all patents according to their common technological relevances. The PDP also assigns an IPC code to each patent using the original USPC and a USPC-IPC concordance based on the International Patent Classification Eighth Edition.

technology class to another are observable. Using the information, the metric (1) is employed to gauge the distance between a new patent and a firm's location in the technology spectrum.

3.2 Stylized Facts

Next, the empirical findings highlighted in the introduction of the paper are presented.

3.2.1 Firm Moments and Patent-Firm Distance

Are patent-firm distances important when it comes to the relationship between a firm's patent portfolio and its moments, such as sales and assets? In order to answer this question, the following table regresses "log real sales" and "log market value" in year t on a firm's patent portfolio, its distance-adjusted patent portfolio, and the firm's size in the same year. To convert nominal variables into their real counterparts, sector-specific NBER CES MID shipment deflators are used. The regressions also include year and firm fixed effects to rule out firm-specific properties and time trends.

Table 1: Firm Moment Regressions

Variable	*log real sales*	*log market value*
log patent stock	0.195***	0.039***
	(0.008)	(0.008)
log dist-adj pat stock	−0.009***	−0.020***
	(0.003)	(0.003)
log employment	0.936***	0.728***
	(0.008)	(0.008)
intercept	yes	yes
year	yes	yes
firm fixed effect	yes	yes
Obs.	23,028	36,094
R^2	0.96	0.92

Standard errors are reported in parentheses.

*10%, **5%, ***1% significance.

As expected, the patent portfolio of a firm is positively related to its sales and stock market valuation. More interestingly, a firm's patent portfolio is negatively related to firm moments, once it is adjusted by patent distances. The distance-adjusted patent stock is constructed in a way such that each patent's contribution to the portfolio is multiplied by its distance to the firm prior to the aggregation. As a result, the coefficient of the distance-adjusted patent stock quantifies the loss of correlation between the patent portfolio and the firm moments due to the technological mismatch between the firm and its patents.

In order to interpret the results correctly, one should consider the ratio of the negative coefficient of the distance-adjusted patent stock to that of the unadjusted patent stock. This ratio is statistically significant, but not huge for real sales (4.62%), but quite high for market value of equity (51.3%). These regressions suggest that a patent portfolio that is perfectly mismatched (with all patents having distance = 1) contributes 50% less to the average firm value compared to a perfectly matched patent portfolio (with all patents having distance = 0). Using the sample means for log patent stock and log distance adjusted patent stock, the percentage loss generated by the observed misallocation turns out to be 2.75% for real sales and 30.5% for market value on average.

3.2.2 Patent Sale Decision and Patent-Firm Distance

Does the technological distance of a patent to the firm influence the decision to keep or sell a patent? In order to conduct this analysis, the indicator variable of whether a patent is sold or kept (=1 if a patent is sold, =0 if not) is regressed on a number of potentially related regressors, including the patent's distance to the initial owner. The following table reports the OLS regression results:

Table 2: Patent Resale Decision

Variable	*Indicator (=1 if sold)*
distance	0.0212***
	(0.001)
patent quality	0.0004***
	(0.000)
log (size of patent portfolio)	−0.0161***
	(0.000)
intercept	yes
year	yes
firm fixed effect	yes
Obs.	2,564,303
R^2	0.4198

Standard errors are reported in parentheses.

*10%, **5%, ***1% significance.

This regression indicates that a patent is more likely to be sold if it is more distant to the firm. The size of the patent portfolio of the firm, and patent quality are also included in the regression to avoid spurious relationships, or mechanical bias that may be caused by the construction of the distance metric. Year and firm fixed effects are likewise included to remove potential sources of bias. Despite the addition of these controls, the coefficient on the distance variable remains statistically significant, and positive. Considering the average number of patents sold ($\approx$ 15%) in the time period, the coefficient suggests that a perfectly mismatched patent is 14.1% ($\approx 0.0212/0.15$) more likely to be sold to another firm, rather than being kept. Recall also that the definition of sale employed is quite conservative, in the sense that patent transfers due to mergers and acquisitions are not considered sales, even though the primary motive for these events might be the acquisition of patents. The results are in line with the intuition that a firm is more likely to sell patents that are not a good fit, rather than keeping them, due to the potential gains from trading the patent to a firm that might be better suited to exploit the embedded ideas commercially.

3.2.3 Patent-Firm Distance Reduction Conditional on a Patent Resale

The primary motivation behind considering patent distance as a likely determinant of patent resale decisions is the potential gains from trade that arise if the patent can be sold to a firm that can use it better, which in expectation yields more profits. If this intuition is correct, the distance between the owner firm and the patent is expected to decrease after a patent is sold. In order to test this hypothesis, a distance difference variable for each patent transaction is constructed, for which the buyer and seller firms can be identified in the PDP data. Here $d(p, f_b)$ denotes the distance of the patent to the buyer firm, and $d(p, f_s)$ to the seller firm. Next, a regression is run on the control variables that consist of year and seller firm fixed effects. The sign and significance of the intercept term is used to test the hypothesis.

Table 3: Distance Reduction on Resale

Variable	Change in distance $d(p, f_b) - d(p, f_s)$
intercept	-0.182^{***}
	(0.061)
year fixed effect	yes
seller fixed effect	yes
Obs.	25,170
R^2	0.3980

Standard errors are reported in parentheses.

*10%, **5%, ***1% significance.

The results indicate that after controlling for year and firm fixed effects, conditional on a patent resale, the distance between a patent and its owner is significantly decreased. In other words, the mismatch between the idea and the firm owning it is reduced. This effect is economically big. Considering that the average measure for distance is 0.56, average reduction in distance is approximately 37.8% ($\approx 0.182/0.481$) of the average distance.

4 Calibration

In order to simulate the model values must be assigned to the various parameters. There are twelve parameters to pick: β, ε, κ, λ, δ, σ, γ, χ, ρ, μ, η, and ω. A distribution needs to be specified for $X(x)$ as well. In order to select values for the parameters a set of data targets is specified for the model to match. Computing the solution to the model essentially involves solving a system of nonlinear equations, as Point (10) in Proposition 2 made clear. A few features of the model, such as the invariant firm-size distribution and the distribution governing the duration of patent sales, are uncovered by conducting a Monte Carlo simulation. In particular, a panel of 100,000 firms is simulated for 3,000 periods. While it is not necessarily the case that a particular parameter governs the ability of model to mimic a certain data target, due to simultaneity in the system of equations that characterizes the model, some intuition about where each parameter is likely to impinge is provided. The data targets are as follows:

1. *Long-run growth in output.* In the U.S. output grew at about 2% over the postwar period. Intuitively, the parameter γ, which enters the law of motion for productivity growth (4), should play an important role in determining this.

2. *Long-run interest rate.* A reasonable value for the long-run interest rate in the U.S. is 4%. Pick β using the equation $\beta = r\mathbf{g}^{\varepsilon\zeta/(\zeta+\lambda)}$– see (24).

3. *The ratio of R&D expenditure to GDP.* The U.S. expenditure on research and development was about 2.91% of GDP. The parameter χ is key here because it governs the cost of doing R&D, as can be seen from (3).

4. *Capital's and labor's shares of income.* In line with Corrado, Hulten and Sichel (2009), capital's (κ) and labor's (λ) share of incomes are selected to 25 and 60%. Intangible capital (ζ) then accounts for the remaining 15%. In the current setting, this is interpreted as the value of patents, or intellectual property, in production.

5. *Depreciation rate for capital.* The depreciation rate of capital is chosen to be 6.9%. This is consistent with the U.S. National Income and Product Accounts.

6. *Survival rate for a patent.* In the U.S. a patent lasts for 17 years. Hence, $\sigma = 1 - 1/(1+17)$.

7. *Fraction of patents sold.* About 16% of patents are sold in U.S. The parameters governing the matching function and the strength of the weight of the patent agent in bargaining underpin this statistic for the model. The higher the weight of the patent agent in the bargaining (ω) the higher will be the price q. This will entice more firms to sell. Similarly, the easier it is to sell a patent, as determined by the matching function, the higher will be the price q.

8. *Duration until a sale.* The empirical frequency distribution for the duration of a sale is targeted. In particular, the calibration procedure tries to minimize the sum of the squared differences between the empirical distribution and the analogue for the model. It takes about 5.34 years on average to sell a patent. The coefficient of variation around this mean is 0.82. So, there is considerable variation in sales duration. The parameters governing the matching function are obviously central here.

9. *The Empirical Distribution for the Proximity of Patent to a Firm's Technology Class.* Empirical distance distributions for the U.S. displayed in Figure 3. For the baseline calibration, take the distance distribution associated with $\iota = 2/3$. Define a measure of closeness or proximity between a patent p and a firm f by $c_\iota(p,f) \equiv 1 - d_\iota(p,f)$, where $d_\iota(p,f)$ is given by (1).The density associated with $c_\iota(p,f)$ is used for $X(x)$. This amounts to just a simple change in units on the horizontal axis in Figure 3. Assume that x is distributed uniformly within each of the ten bins of the histogram.

10. *R&D Cost Elasticity.* In order to estimate the elasticity of the R&D cost function, the cost function in the model is inverted to obtain a production function. Then a

regression is run using Compustat data to determine the parameter values, where the output of the R&D production function is proxied for by citation-weighted patents.

11. *CRRA parameter.* This parameter is taken to be 2, the midpoint between varying estimates reported in Kaplow (2005).

12. *Bargaining power.* The bargaining powers of buyers and sellers are chosen to be equal.

The upshot of the calibration procedure is displayed in Tables 4 and 5. The parameters β, ε, κ, λ, δ, σ, γ, ρ, and ω can be calculated directly from the data targets (2), (4), (5), (6), (10), (11) and (12) without having to solve the model. The distribution $X(x)$ comes from (9). To obtain values for χ, μ, and η the solution for the model must be computed. The values for these parameters are selected by minimizing the Euclidean distance between the 5 data targets outlined in (1), (3), (7) and (8) and the model's predictions for these targets.

TABLE 4: PARAMETER VALUES

Parameter Value	*Description*	*Identification*
$\beta = 0.98$	Discount factor	Real interest rate
$\varepsilon = 2.00$	CRRA parameter	Kaplow (2005)
$\kappa = 0.25$	Capital's share	Corrado, Hulten and Sichel (2006)
$\lambda = 0.60$	Labor's share	"
$\delta = 0.07$	Depreciation rate	NIPA
$\sigma = 0.94$	Patent survival rate	U.S. patent law
$\gamma = 0.38$	Law of motion, productivity	Growth rate in GDP
$\chi = 2.49$	Cost of R&D	R&D expenditure to GDP
$\rho = 3.00$	R&D cost elasticity	Regression using Compustat data
$\mu = 0.52$	Matching function	Fraction of patents sold, duration
$\eta = 0.15$	"	until sale, and c.v. over duration
$\omega = 0.50$	Bargaining power	Imposed, equal for buyers and sellers
$X(x)$	Proximity distribution	Empirical distribution

TABLE 5: CALIBRATION TARGETS

Target	*U.S. Data*	*Model*
Long-run growth in output	2.00%	2.00%
Ratio of R&D expenditure to GDP	2.91%	2.41%
Fraction of patents that are sold	15.6%	15.2%
Average duration until a sale	5.34 yrs.	5.28 yrs.
Sale duration, c.v	0.82	1.09

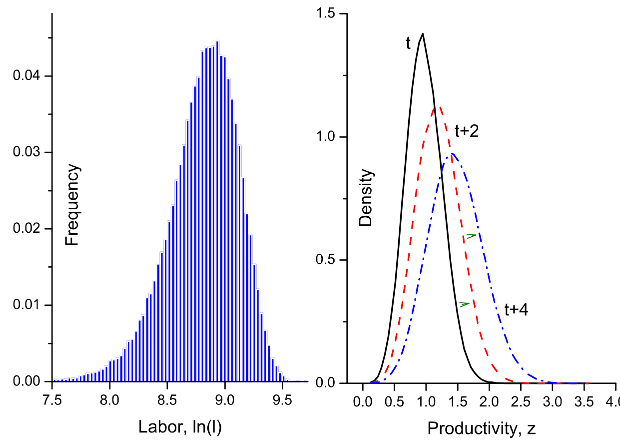

Figure 5: The distributions for firm size (left panel) and productivity (right panel).

The unique invariant firm-size distribution associated with this calibration is shown by the left panel of Figure 5. This distribution resembles a log normal. The coefficient of variation for the distribution is 50%. This falls short of the 200% observed in the data. So, it should though. Surely all differences in productivity across firms are not accounted for simply by differences in innovation. The right panel of Figure 5 illustrates how the productivity distribution shifts rightward over time due to growth in the economy. (Here the histograms calculated from the Monte Carlo are replaced by a fitted density function so the movement could be highlighted). A change in shape of the distribution over time is evident. Last, Figure 6 shows, for both the data and model, the frequency distribution over the duration for a sale. As can be seen, it appears to be harder to affect a sale in data than in the model.

5 Findings

The importance of a resale market for patents will be gauged. To do this, various experiments that change the efficiency of the patent market will be entertained. The efficiency of the

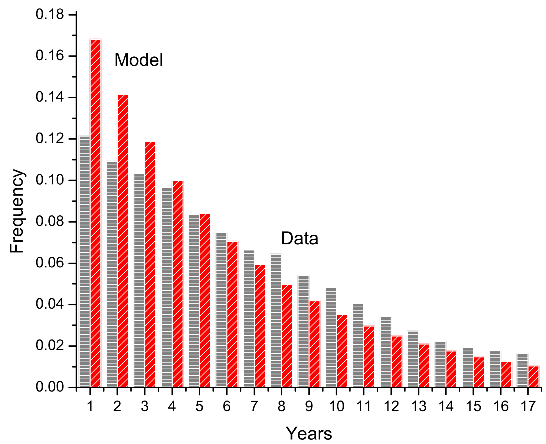

Figure 6: Sale duration distributions, data and model.

resale market for patents is increased in stages. First, an experiment is performed where the meeting rate for matches is allowed to rise. While it may be easier for buyers and sellers to meet now, a seller's idea may still not be well suited for the buyer. The next experiment considers a situation where patent agents can find buyers who are perfect matches for the ideas that they are selling. So, there is no mismatch between buyers and sellers on the patent market. Still, innovating firms produce ideas that are not ideally suited for their own businesses and this injects a friction into the analysis. A patent that is not incorporated into an innovator's production process will only have a finite life on the resale market. Additionally, it may take time to find a buyer. The final experiment focuses on the case where innovating firms produce ideas that are tailored toward their own production activity. Here ideas are perfectly matched with the developer. The change in welfare from moving from one environment to another is calculated. The metric for comparing welfare will be discussed now.

5.1 Welfare Comparisons

Consider two economies, namely A and B, moving along their balanced growth paths. The consumption/output ratio and gross growth rate for economy A are given by $\mathbf{c}^A/\mathbf{o}^A$ and $\mathbf{g}^A$.

Similar notation is used for country B. To render things comparable, start each country off from an initial position where $\mathbf{o}^A = \mathbf{o}^B = 1$. Thus, $\mathbf{c}^A = \mathbf{c}^A/\mathbf{o}^A$ and $\mathbf{c}^B = \mathbf{c}^B/\mathbf{o}^B$; i.e., the initial consumption/output ratio for each country can be used as measure of initial consumption. Now, the levels of welfare for economies A and B are given by

$$W^A = \sum_{t=1}^{\infty} \beta^{t-1} \frac{(\mathbf{c}_t^A)^{1-\varepsilon}}{1-\varepsilon} = \frac{(\mathbf{c}_1^A/\mathbf{o}_1^A)^{1-\varepsilon}}{(1-\varepsilon)[1-\beta(\mathbf{g}^A)^{1-\varepsilon}]},$$

and

$$W^B = \frac{(\mathbf{c}_1^B/\mathbf{o}_1^B)^{1-\varepsilon}}{(1-\varepsilon)[1-\beta(\mathbf{g}^B)^{1-\varepsilon}]}.$$

How much would initial consumption in economy A have to be raised or lowered to make people have the same welfare level as in economy B? Denote the fractional amount in gross terms by α (which may be less than one). Then, α must solve

$$\frac{(\alpha\mathbf{c}_1^A/\mathbf{o}_1^A)^{1-\varepsilon}}{(1-\varepsilon)[1-\beta(\mathbf{g}^A)^{1-\varepsilon}]} = W^B$$

so that

$$\alpha = (W^B/W^A)^{1/(1-\varepsilon)}.$$

This will be the welfare measure that is used in all experiments.

5.2 Increasing the Contact Rate for Matches, η

The patent resale market mitigates the initial misallocation of ideas. Still, it takes time to sell a patent as the patent agent may not be able to find a buyer. To understand how this friction in matching affects the economy, it is useful to look at how a change in the scale factor for the matching function, η, affects key variables. Figure 7 summarizes the results.

A rise in the contact rate reduces the time that it takes to find a buyer, as the lower panel of Figure 7 illustrates. The rate of innovation also falls, which is more surprising. The consequences of failing to innovate are now lessened because it will be easier for a firm to buy a patent. Growth (upper panel) increases along with efficiency in matching. So does welfare. On the one hand, if patent resale market was completely closed, the welfare loss

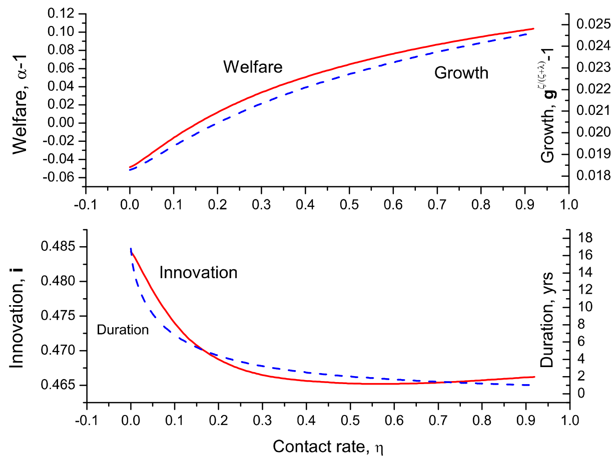

Figure 7: The impact of an increase in contact rate for matching on sales duration, growth, innovation and welfare.

would be equivalent to losing approximately 4.8% of initial consumption. On the other hand, if the efficiency of the market was at its extreme (the minimum value for η that results in all buyers meeting a patent agent with probability 1), welfare would be 10.4% higher than the calibrated economy. The upshot is that the resale market for patents plays an important role in the economy.

5.3 Perfectly Directed Search

A second source of inefficiency in the model is the random search technology used in the patent market. In the baseline model, conditional upon a meeting between a buyer and a resale agent, the proximity of the patent to the firm is drawn from a uniform distribution. Instead imagine a directed search structure, where patent agents are able to target the segment of the economy that exactly matches the patent they want to sell. In such a case, whether or not a patent agent meets a buyer is still a probabilistic event governed by the matching function. The proximity between the patent and the firm would be nonstochastic and equal to unity; in other words, a perfect match. The level of welfare in this alternative economy is 7.4% higher than in the baseline one. The output growth rate increases slightly from 2% to 2.3%, despite a small decline in the innovation rate. The fraction of patents sold increases from 15.2 to 24.2%. But most strikingly, the decomposition of growth reveals that fraction of growth due to patents sold doubles from 15.7 to 34.8%. This decomposition can be done using equation (27). Note that there are two terms in the brackets. The first term can be used to measure the contribution to growth from the ideas that firms keep, the second from the ones that they sell. The table below summarizes the results.

Table 6: Perfectly Directed Search

	Baseline Model	*Directed Search*
Output growth rate, $\mathbf{g}^{\zeta/(\zeta+\lambda)} - 1$	0.020	0.023
Innovation rate, $\mathbf{i}$	0.47	0.45
Welfare gain, $\alpha - 1$	0.00	0.07
Fraction of patents sold	0.15	0.24
Growth from patents sold	0.16	0.35

5.3.1 Perfectly Directed Search with a High Contact Rate

Now, redo the above experiment while also using a high contact rate for matches. Output growth is now much higher at 3.2%, even though innovation is slightly lower than in the baseline model. This reflects a reduction in misallocation. The reduction in innovation is reflected by a slightly higher consumption/output ratio. As can be seen, now most patents are sold. Economic welfare is 28% higher.

Figure 8 gives the upshot from the experiments that have been conducted so far. It shows how the cumulative distribution function for ideas, or for x, changes across the various experiments. First, in the data firms produce ideas that are not well suited for their own lines of business, as can be seen from the distribution labeled "Empirical". (Recall that a higher value for $x \in [0, 1]$ indicates that an idea is better suited for the firm's business activity.) In the baseline model, a firm is free to sell such an idea. A firm that fails to innovate is free to buy one from another firm. This leads to a better distribution of ideas, as is reflected in the distribution function for the baseline model after transactions on the secondary market for patents have been consummated. The distribution function for the baseline model stochastically dominates, in the first-order sense, the empirical distribution. When the contact rate

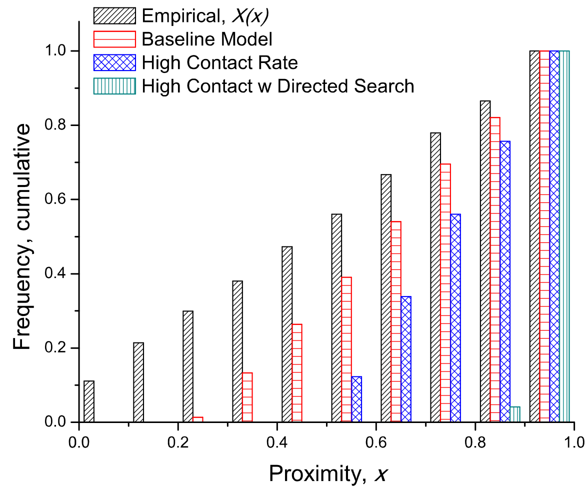

Figure 8: Misallocation of ideas. Cumulative distribution functions for x.

for matching is high it relatively easy to consummate a patent sale. The distribution for x improves–see the histogram labeled "High Contact Rate", which stochastically dominates the one for the baseline model. Of course, if search could be perfectly directed things would be better still–"High Contact w Directed Search", which stochastically dominates all other distributions.

Note that not all firms sell their patents, even though they are not perfectly matched with their ideas. This occurs because are still some frictions left in the patent market. First, there are more sellers than buyers on the market, so not all patents will be immediately sold. Second, patents have a finite life on the market and hence suffer some depreciation. Both these factors imply that the price at which a firm can sell a patent, q, will be less than what it is worth to a perfectly matched firm.

Table 7: Perfectly Directed Search with High Contact

	Baseline Model	*Directed Search with High Contact*
Output growth rate, $\mathbf{g}^{\zeta/(\zeta+\lambda)}-1$	0.020	0.031
Innovation rate, $\mathbf{i}$	0.47	0.468
Welfare gain, $\alpha-1$	0.00	0.28
Fraction of patents sold	0.15	0.81
Growth from patents sold	0.16	0.83
Consumption/output	0.805	0.813
Duration, yrs	5.23	0.00
Seller/buyers	1.23	0.69
mean, x–after patent transactions	0.67	0.99
St. dev, x–after patent transactions	0.21	0.03

5.4 Removing the Misallocation of Ideas

The central inefficiency in the framework derives from the fact that firms develop ideas that are imperfect matches for the own production processes. The presence of a secondary market for patents mitigates this problem. Suppose that an innovating firm always comes up with an idea that is a perfect match for its production process. That is, let each innovating firm always draw $x=1$. In this situation, the economy could increase its growth rate from 2 to 3.4%, a big jump. Welfare would increase by 35%. This illustrates that the frictions arising from mismatches in innovation are large.

Table 8: Perfect Innovation

	Baseline Model	*Perfect Innovation*
Output growth rate, $\mathbf{g}^{\zeta/(\zeta+\lambda)}-1$	0.020	0.034
Innovation rate, $\mathbf{i}$	0.47	0.49
Welfare gain, $\alpha-1$	0.00	0.35

5.5 Importance of the Seller's Bargaining Power, ω

The baseline calibration assumes that the bargaining power of buyers and sellers (patent agents) are equal. It turns out that this gives the best fit for calibration. It's instructive to show how the model behaves for varying values of ω, the bargaining power of patent agents. To see this, direct attention to Figure 9.

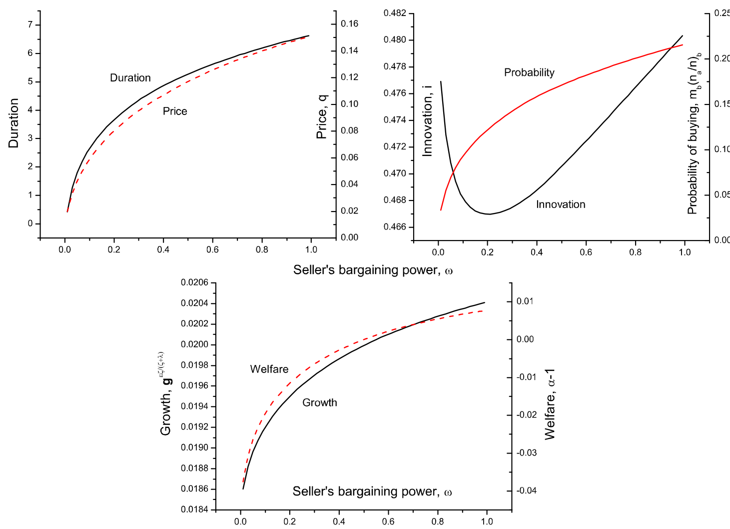

Figure 9: The impact of the seller's bargaining power on sales duration, the price of a patent, innovation, the odds of finding a seller, growth and welfare.

To begin with, observe that the price received by the selling firm, q, rises continuously with the bargaining power of the patent agent, ω. Recall that the selling firm takes all of the surplus earned by the patent agent, as reflected in the form of (20). Next, as the price rises so does the average time it takes to affect a sale. This makes sense. Buyers will become choosier about the patents they buy as prices rise; i.e., they will demand patents that are closer to their line of business. Interestingly, the rate of innovation is $\cup$ shaped in ω. There are two forces at work here. On the one hand, prices rise with the seller's bargaining power. This increases the value of a selling firm, $S(z; \mathbf{s})$, and spurs innovation, as can be seen from (14). On the other hand, the odds of buying a patent increase with ω, as there are more sellers on the market relative to buyers. The value of a buying firm, $B(z; \mathbf{s})$, rises on this account. This lowers the incentive to innovate; again see (14). At low values for ω the second effect dominates while at high values the first one does. This results in the $\cup$ shape. Despite the nonmonotone behavior for innovation, both growth and welfare increase in ω. It is better for the economy when innovating firms capture the payoffs from developing new ideas.

6 Robustness Analysis

Central to the analysis is the notion of how distant a new patent is to a firm's location in the technology spectrum. The distance between a new patent and a firm's patent portfolio is measured in a fairly general way using the metric (1), which is governed by the parameter ι. Figure 3 illustrates how the empirical distance distribution changes with ι. When $\iota = 1$ the distance metric just computes the average distance between the new patent and the other patents in the firm's portfolio. Lower values for ι put higher weight on patents in the portfolio that are closer to the new patent. The analysis in the paper is done for three values of ι : $\iota = 1/3$. $\iota = 2/3$ (the baseline model), and $\iota = 1$ (average distance). The results hold for all three values of ι. Appendix 9.3 details some of the findings. In a nutshell:

1. The regression results reported in Section 3 do not change in a material manner.

2. The calibration targets can be hit for all three values of ι.

3. The model behaves essentially the same way for the three distance measures. For example, shutting down the patent resale market in the baseline model led to a welfare loss of 4.8% in terms of consumption. This welfare loss rose to 5.4%, when $\iota = 1$, and fell to 3.5%, when $\iota = 1/3$. Lower values for ι compress the distribution for new ideas; i.e., new ideas are closer on average to the firm's location on the technology spectrum. Therefore, shutting down the patent market has a smaller effect when ι is lower.

The empirical sale duration distribution is another object that disciplines the model. In the baseline, the duration is defined as the time between the sale and grant dates of all patents sold between 1980 and 2012. Different methodologies can be considered. One can consider using the application date instead of the grant date since some patents are sold before they are granted. Patents need to be observed by potential buyers before being considered "on the market". The date of the first citation a patent receives can be a good proxy for this. Lastly, it is possible that excluding more recent observations might prevent the confounding effects of a potential truncation bias. The analysis is repeated with the three new sale distributions, and the main findings remain robust. Appendix 9.3 details some of the findings.

7 Conclusions

A model of the market for patents is developed here. Each period a firm conducts research and development. This R&D process may spawn ideas. Some of the ideas are useful for a firm's line of business, others are not. A firm can patent and sell the ideas that are not. The fact it can sell ideas provides an incentive to engage in R&D. Likewise, firms that fail to innovate can attempt to buy ideas. This allows a firm to grow its business. The efficiency of the patent market for matching ideas with firms has implications for growth. These are examined here.

Empirical analysis, drawing on the NBER-USPTO patent grant database and patent reassignment data available from Google Patent Beta, established five useful facts. First, somewhere between 15 and 20% of patents are sold. Second, it takes on average 5.34 years to sell a patent. Third, a firm's patent stock contributes more to its market value and sales the closer it is to the firm in terms of average technological distance. Fourth, a patent is more likely to be sold the more distant it is to a firm's line of business. Fifth, when a patent is sold it is closer to the buyer's line of business than to the seller's. These facts suggest that a resale market for patents may play a important role in correcting the misallocation of ideas across firms. It may also influence a firm's R&D decision.

The developed model is calibrated to match several stylized facts characterizing the U.S. data, such as the postwar rate of growth, the ratio of R&D spending to GDP, the fraction of patents sold, the average time it takes to sell a patent, and the dispersion in the length of time that it takes to sell a patent. The importance of a market for selling patents is then assessed. This is done by conducting a series of thought experiments that successively increase the efficiency of the patent market. The efficiency of this market is important for economic growth and welfare.

References

Acemoglu, Daron, Philippe Aghion and Fabrizio Zilibotti (2006). "Distance to Frontier, Selection, and Economic Growth." *Journal of the European Economic Association*, 4(1): 37–74.

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr (2013)."Innovation, Reallocation and Growth." Unpublished paper, University of Pennsylvania.

Aghion, Philippe, Ufuk Akcigit and Peter Howitt (2013)."What Do We Learn From Schumpeterian Growth Theory?" NBER Working Paper 18824.

Benhabib, Jess, Jesse Perla and Christopher Tonetti (2012). "Catch-up and Fall-back through Innovation and Imitation." NBER Working Paper 18091.

Chatterjee, Satyajit and Esteban Rossi-Hansberg (2012). "Spinoffs and the Market for Ideas." *International Economic Review* 53(1): 53-93.

Chiu, Jonathan, Cesaire Meh and Randall Wright (2011). "Innovation and Growth with Financial, and Other, Frictions." NBER Working Paper 17512.

Corrado Carol, Charles R. Hulten and Daniel E. Sichel (2009). "Intangible Capital and U.S. Economic Growth." *The Review of Income and Wealth* 55(3): 661-685.

Gans, Joshua S. and Scott Stern (2010). "Is there a Market for Ideas?" *Industrial and Corporate Change*, 19(3): 805–837.

Guner, Nezih , Gustavo Ventura, and Xu Yi (2008). "Macroeconomic Implications of Size-Dependent Policies." *Review of Economic Dynamics*, 11(4): 721-744.

Hagiu, Andrei and David Yoffie (2011). "Intermediaries for the IP Market." Working Paper 12-023, Harvard Business School.

Hall, Bronwyn H., Adam Jaffe and Manuel Trajtenberg (2005). "Market Value and Patent Citations." *Rand Journal of Economics*, 36(1): 16-38.

Hsieh, Chang-Tai and Peter J. Klenow (2009). "Misallocation and Manufacturing: TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403-1448.

Jovanovic, Boyan and Glenn M. MacDonald (1994). "The Life Cycle of a Competitive Industry." *Journal of Political Economy*, 102(2): 322-347.

Kaplow, Louis (2005). "The Value of a Statistical Life and the Coefficient of Relative Risk Aversion." *Journal of Risk and Uncertainty* 31(1): 23-34.

König, Michael, Jan Lorenz and Fabrizio Zilibotti (2012). "Innovation vs Imitation and the Evolution of Productivity Distributions," CEPR Discussion Paper 8843.

Lamoreaux, Naomi R. and Kenneth L. Sokoloff (2003). “Intermediaries in the U.S. Market for Technologies, 1870-1920.” In Stanley Engerman et al (Editors) *Finance, Intermediaries, and Economic Development*, Cambridge University Press.

Lucas, Robert E. Jr. and Benjamin Moll (2011). “Knowledge Growth and the Allocation of Time.” NBER Working Paper 17495.

Restuccia, Diego and Richard Rogerson (2008). “Policy Distortions and Aggregate Productivity with Heterogeneous Plants.” *Review of Economic Dynamics*, 11(4): 707-720.

Serrano, Carlos J. (2010). “The dynamics of the transfer and renewal of patents.” *RAND Journal of Economics* 41(4): 686-708.

Serrano, Carlos J. (2011). “Estimating the Gains from Trade in the Market for Innovation: Evidence from the Transfer of Patents.” NBER Working Paper No. 17304.

Stokey, Nancy L. and Robert E. Lucas, Jr. with Edward C. Prescott (1989). *Recursive Methods in Economic Dynamics*. Cambridge, MA: Harvard University Press.

8 Theory Appendix

8.1 Balanced Growth

Proof. The proof proceeds using a guess and verify procedure (or the method of undetermined coefficients).

Point (1). To derive the interest and rental rates, imagine the problem of a consumer/worker who can invest in one period bonds that pay a gross interest rate of $1/r$. The Euler equation for asset accumulation will read

$$c^{-\varepsilon} = (\beta/r)(c')^{-\varepsilon}.$$

Along a balanced growth path, if the mean level of productivity grows at rate $\mathbf{g}$ then consumption, the capital stock and output must grow at rate $\mathbf{g}^{\zeta/(\zeta+\lambda)}$. This fact can be gleaned from the production function (2), by assuming z grows at rate $\mathbf{g}$, that capital and output grow at another common rate, and that labor remains constant. Therefore, $r = \beta/\mathbf{g}^{\varepsilon\zeta/(\zeta+\lambda)}$. In standard fashion, the rental rate on capital is given by $\widetilde{r} = 1/r - 1 + \delta = \mathbf{g}^{\varepsilon\zeta/(\zeta+\lambda)}/\beta - 1 + \delta$.

Point (4). The form of the threshold rule for buying a patent follows from the fact the sum of the surplus (sans price) accruing to a firm that buys a patent and the surplus (sans price) to the patent agent must be greater than zero; otherwise, a non-negative sale price, p, for the patent would not exist. First, plug the solutions for w and $\widetilde{r}$, or (19) and (25), into the profit function (7) to obtain

$$\Pi(z, \mathbf{s}) = \pi \frac{z}{\mathbf{z}^{\lambda/(\zeta+\lambda)}} = \pi \tilde{z},$$

with

$$\pi \equiv \frac{\zeta}{\mathbf{g}^{\lambda/(\zeta+\lambda)}} \left(\frac{\kappa}{\mathbf{g}^{\zeta/(\zeta+\lambda)}/\beta + \delta - 1} \right)^{[\kappa - \lambda\kappa/(\zeta+\lambda)]/\zeta}. \qquad (32)$$

Second, conjecture that the value functions $V(z;\mathbf{s})$ and $A(\mathbf{s})$ have the forms $V(z;\mathbf{s}) = \mathfrak{v}_1 \widetilde{z} + \mathfrak{v}_2 \widetilde{\mathbf{z}}$ and $A(\mathbf{s}) = \mathfrak{a}\widetilde{\mathbf{z}}$. Third, given the above, note that the (sans price) surpluses for a buying firm and the patent agent are given by

$$\pi(\tilde{z} + \gamma x \widetilde{\mathbf{z}}) - \pi \tilde{z} + rV(z + \gamma x \mathbf{z}, \mathbf{s}') - rV(z, \mathbf{s}') = (\pi + \frac{r\mathfrak{v}_1}{\mathbf{g}^{\lambda/(\zeta+\lambda)}})\gamma x \widetilde{\mathbf{z}},$$

and

$$-\sigma r A(\mathbf{s}') = -\sigma r \mathbf{g}^{\zeta/(\zeta+\lambda)} \mathfrak{a} \widetilde{\mathbf{z}} \text{ [cf. (17)].}$$

Finally, note that whether or not the sum of the above two equations is nonnegative does not depend on $\widetilde{\mathbf{z}}$. This sum is also increasing in x. Solving for the value of x that sets the sum to zero yields

$$x_a = \frac{\sigma r \mathbf{g}^{\zeta/(\zeta+\lambda)} \mathfrak{a}}{(\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\gamma}. \tag{33}$$

Thus, x_a is a constant.

Point (8). The solutions for patent prices are easy to obtain. Insert the above formulae for the (sans price) surplus for a buying firm and the (sans price) surplus for a patent agent into expression (16) to get

$$P(\mathbf{z}, x; \mathbf{s}) = \left[(1-\omega)\sigma r \mathbf{g}^{\zeta/(\zeta+\lambda)} \mathfrak{a} + \omega(\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\gamma x\right] \widetilde{\mathbf{z}}.$$

It is immediate from (20) that $q = \mathfrak{a}\widetilde{\mathbf{z}}$, predicated upon the guess $A(\mathbf{z}) = \mathfrak{a}\widetilde{\mathbf{z}}$.

Point (5). It will now be shown that the value function for the patent agent has the conjectured linear form. Focus on equation (15), which specifies the solution for A. The price for a patent does not depend on z, given Point (8). Additionally, $D(x) = \mathcal{U}[0,1]$. Furthermore, $I_a = 1$ for $x > x_a$ and is zero otherwise. Thus,

$$A(\mathbf{s}) = \mathfrak{a}\widetilde{\mathbf{z}} = m_a(n_a/n_b) \int_{x_a}^{1} P(\mathbf{z}, x; \mathbf{s}) dx + [1 - m_a(n_a/n_b) Pr(x \geq x_a)] \sigma r A(\mathbf{s}'),$$

from which it follows that

$$\begin{aligned} \mathfrak{a} &= \sigma r \mathbf{g}^{\zeta/(\zeta+\lambda)} \mathfrak{a} - m_s(n_a/n_b)(1-x_a)\omega\sigma r \mathbf{g}^{\zeta/(\zeta+\lambda)} \mathfrak{a} \\ &\quad + m_s(n_a/n_b)\omega(\pi + r\mathfrak{v}_1/\mathbf{g}^{\zeta/(\zeta+\lambda)})\gamma(1-x_a)(1+x_a)/2. \end{aligned} \tag{34}$$

Point (2). The value function for a buying firm can be determined in a manner similar to that for A in Point (5). Here

$$B(z, \mathbf{z}) = \mathfrak{b}_1 \tilde{z} + \mathfrak{b}_2 \widetilde{\mathbf{z}},$$

with

$$\mathfrak{b}_1 = \pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)}, \tag{35}$$

and

$$\mathfrak{b}_2 = -m_b(n_a/n_b)(1-x_a)(1-\omega)\sigma r \mathbf{g}^{\zeta/(\zeta+\lambda)}\mathfrak{a} + r\mathfrak{v}_2\mathbf{g}^{\zeta/(\zeta+\lambda)} \quad (36)$$
$$+m_b(n_a/n_b)(1-\omega)(\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\gamma(1-x_a)(1+x_a)/2.$$

To derive this solution, the results in Points (4) and (8), along with the conjectured solution for V, are used in equation (8). Similarly, using equation (11) it can be shown that the value function for a seller is given by

$$S(z,\mathbf{s}) = \mathfrak{s}_1\tilde{z} + \mathfrak{s}_2\widetilde{\mathbf{z}},$$

with

$$\mathfrak{s}_1 = \pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)}, \quad (37)$$

and

$$\mathfrak{s}_2 = \sigma\mathfrak{a} + r\mathfrak{v}_2/\mathbf{g}^{\zeta/(\zeta+\lambda)}. \quad (38)$$

Last, following from (10),

$$K(L(z,x;\mathbf{s});\mathbf{s}) = \mathfrak{k}_1\tilde{z} + \mathfrak{k}_2(x)\widetilde{\mathbf{z}},$$

with

$$\mathfrak{k}_1 = \pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)}, \quad (39)$$

and

$$\mathfrak{k}_2(x) = (\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\gamma x + r\mathfrak{v}_2/\mathbf{g}^{\zeta/(\zeta+\lambda)}. \quad (40)$$

Point (3). The threshold rule for keeping or selling a patent is determined by the condition

$$\mathfrak{k}_1\tilde{z} + \mathfrak{k}_2(x_k)\widetilde{\mathbf{z}} = \mathfrak{s}_1\tilde{z} + \mathfrak{s}_2\widetilde{\mathbf{z}};$$

that is, at the threshold a firm is indifferent between keeping or selling the patent. Now, $\mathfrak{s}_1 = \mathfrak{k}_1$ so that

$$(\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\gamma x_k + r\mathfrak{v}_2\mathbf{g}^{\zeta/(\zeta+\lambda)} = \sigma\mathfrak{a} + r\mathfrak{v}_2\mathbf{g}^{\zeta/(\zeta+\lambda)}.$$

Hence,

$$x_k = \frac{\sigma\mathfrak{a}}{[\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)}]\,\gamma}, \quad (41)$$

a constant.

Point (6). Turn now to the beginning-of-period value function for the firm and the rate of innovation that it will choose. By using the linear forms for the value functions $B(z,\mathbf{z})$, $K(z,\mathbf{z},x)$, and $S(z,\mathbf{z})$, the fact that $\mathfrak{b}_1 = \mathfrak{k}_1 = \mathfrak{s}_1$, and the property that the threshold rule takes the form $I_k = 1$ for $x > x_k$ and $I_k = 0$ otherwise, the firm's dynamic programming problem (13) can be rewritten as

$$V(z,\mathbf{s}) = \widetilde{\mathbf{z}} \max_{i\in[0,1]}\{[X(x_k)\mathfrak{s}_2 + \int_{x_k} \mathfrak{k}_2(x)dX(x) - \mathfrak{b}_2]i - \frac{\chi}{1+\rho}i^{1+\rho}\} + (\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\tilde{z} + \mathfrak{b}_2\widetilde{\mathbf{z}}.$$

Differentiating with respect to i then gives

$$X(x_k)\mathfrak{s}_2 + \int_{x_k} \mathfrak{k}_2(x)dX(x) - \mathfrak{b}_2 = \chi i^{\rho},$$

from which (26) follows. Using the solution for i, as given by (26), in the above programming problem yields

$$V(z,\mathbf{s}) = \frac{\rho}{(1+\rho)\chi^{1/\rho}}[X(x_k)\mathfrak{s}_2 + \int_{x_k} \mathfrak{k}_2(x)dX(x) - \mathfrak{b}_2]^{1+\frac{1}{\rho}}\widetilde{\mathbf{z}} + (\pi + r\mathfrak{v}_1/\mathbf{g}^{\lambda/(\zeta+\lambda)})\tilde{z} + \mathfrak{b}_2\widetilde{\mathbf{z}}.$$

It then follows that

$$\mathfrak{v}_1 = \frac{\mathbf{g}^{\lambda/(\zeta+\lambda)}}{\mathbf{g}^{\lambda/(\zeta+\lambda)} - r}\pi, \tag{42}$$

and

$$\mathfrak{v}_2 = \mathfrak{b}_2 + \frac{\rho}{(1+\rho)\chi^{1/\rho}}[X(x_k)\mathfrak{s}_2 + \int_{x_k} \mathfrak{k}_2(x)dX(x) - \mathfrak{b}_2]^{1+\frac{1}{\rho}}. \tag{43}$$

Point (7). The gross rate of growth for aggregate productivity, $\mathbf{g}$, will now be calculated. Suppose that aggregate productivity is currently $\mathbf{z}$. A fraction $\mathbf{i}$ of firms will innovate today. Those firms that draw $x > x_k$ will keep their patent. The productivity for these firms will increase. The fraction $1 - \mathbf{i}$ of firms will fail to innovate. Out of these firms the proportion $m_b(n_a/n_b)$ will find a seller on the patent market. They will buy a patent when $x > x_a$. Thus, $\mathbf{z}$ will evolve according to

$$\mathbf{z}' = \mathbf{z} + \mathbf{i}\int_{x_k}^{1} \gamma x\mathbf{z}dX(x) + m_b(n_a/n_b)(1-\mathbf{i})\int_{x_a}^{1} \gamma x\mathbf{z}dx.$$

This implies (27).

Point (9). The number of buyers on the patent market is given by $n_b = 1 - \mathbf{i}$; all firms that fail to innovate will try to buy a patent. Along a balanced growth path, the number of patent agents, n_a, must satisfy the equation

$$n_a = \sigma n_a[1 - m_a(n_a/n_b)(1 - x_a)] + \sigma \mathbf{i} X(x_k).$$

Focus on the righthand side. Take the first term. Suppose that there are n_a patent agents at the beginning of the period. A fraction σ of these agents will survive. Out of these, $m_b(n_a/n_b)(1-x_a)$ will find a buyer. Thus, they will not be around these next period. Move to the second term. A mass of $\mathbf{i}X(x_k)$ new firms will decide to sell their patents Out of this σ will survive. The sum of these two terms equals the new stock of patent for sale, n_a. Solving yields

$$n_a = \frac{\sigma \mathbf{i} X(x_k)}{1 - \sigma[1 - m_a(n_a/n_b)(1 - x_a)]} \text{ and } \frac{n_a}{n_b} = \frac{\sigma \mathbf{i} X(x_k)}{(1-\mathbf{i})\{1 - \sigma[1 - m_a(n_a/n_b)(1 - x_a)]\}}.$$

Equations (28) and (29) follow immediately.

Point (10). The 12 constants, viz $\mathfrak{a}$, $\mathfrak{b}_1$, $\mathfrak{b}_2$, $\mathfrak{k}_1$, $\mathfrak{k}_2$, π, $\mathfrak{s}_1$, $\mathfrak{s}_2$, $\mathfrak{v}_1$, $\mathfrak{v}_2$, x_a and x_b, are specified by the 12 non-linear equation (32) to (43). The equations include the variables $\mathbf{g}, \mathbf{i}, r, m_a(n_a/n_b)$, and $m_b(n_a/n_b)$. So, equations (24), (26), (27), (28) and (29) must be appended to the system to obtain a system of 17 equations in 17 unknowns. This system does not depend on either $\widetilde{z}$ or $\widetilde{\mathbf{z}}$. ■

8.2 Existence of a Unique Stationary Firm-Size Distribution

Proof. By Stokey and Lucas (1989, Theorem 12.12), it is sufficient to establish three things. First, the transition operator associated with (31) needs to satisfy the Feller Condition [see Stokey and Lucas (1989, p. 220]. Second, it is required that this transition operator is monotone [Stokey and Lucas (1989, p. 220]. Third, the transition operator must satisfy a mixing condition [Stokey and Lucas (1989, Assumption 12.1)].

The stochastic difference equation (31) is continuous in $\widehat{z}$, trivially. It then follows, using Stokey and Lucas (1989, Theorem 8.9 and Exercise 8.10), that a transition operator connected with (31) exists and satisfies the Feller property. Denote this operator by $P(\widehat{z}, \mathcal{B})$, which gives the probability measure connected with a move for z from the point $\widehat{z}$ into the set $\mathcal{B}$. Similarly, let $Q(\mathcal{B})$ represent the probability measure connected with drawing an $x \in \mathcal{B}$. To establish monotonicity, consider the integral

$$\int H(z')P(\widehat{z}, d\widehat{z}') = \int H(\frac{1}{\mathbf{g}}\widehat{z} + \frac{\gamma}{\mathbf{g}}x)Q(dx),$$

for any non-decreasing function $H(z')$. Clearly,

$$\int H(\frac{1}{\mathbf{g}}\widehat{z} + \frac{\gamma}{\mathbf{g}}x)Q(dx) \geq \int H(\frac{1}{\mathbf{g}}\underline{z} + \frac{\gamma}{\mathbf{g}}x)Q(dx),$$

for all $\widehat{z} > \underline{z}$. Thus, P is monotone.

Finally, turn to the mixing condition. The long-run mean of the above process is $z^* = E[x]\overline{z}$, where $0 < z^* < \overline{z}$. To satisfy the mixing condition, it suffices to show that if the process starts off at $\widehat{z} = 0$ then there exists some chance that it will cross into the interval $[z^*, \overline{z}]$, and analogously if originates at the $\widehat{z} = \overline{z}$ then there are some odds that it will cross into the set $[0, z^*]$. Clearly, if the process starts off from $\widehat{z} = \overline{z}$ then there are some odds that it will cross into the set $[0, z^*]$. The firm can draw 0 with strictly positive probability, as can be seen from (30). So, just think about drawing 0 for some prolonged but finite period of time, $T + 1$. Eventually, the process will cross into $[z^*, \overline{z}]$; this will take a maximum of $T + 1$ periods where $T = \ln(z^*/\overline{z})/\ln(1/g)$. This occurs with probability, $\{(1-i)[(1-m_b) + m_b x_a] + i x_k\}^{T+1} > 0$. Likewise, if the process starts off from $\widehat{z} = 0$ then there are some odds that it will cross into the set $[z^*, \overline{z}]$. Think about drawing an x shock in the interval $[\underline{x}, 1]$, where $\underline{x} = \max\{E[x] + \varepsilon, x_a, x_k\}$ with $\varepsilon > 0$. This occurs with some strictly positive probability denoted here, in this proof only, by ξ–again see (30). Imagine drawing this shock for some long, finite period of time. Then, it will take a maximum of $t+1$ periods for the process to cross into the set $[z^*, \overline{z}]$, where $t = \ln[1 - (z^*/\underline{x})(1-\rho)/\rho]/\ln\rho$, with $\rho \equiv \gamma/\mathbf{g}$. The probability of this occurring is $\xi^{t+1} > 0$. ■

9 Empirical Appendix

9.1 Data Description

In a nutshell, data from three sources are used: United States Patent and Trademark Office (USPTO), NBER Patent Database Project (PDP), and Compustat. The first sources contains data on the patents that are reassigned across firms. The second is used to retrieve information on the technological classes for patents and the stocks of patents for firms. Facts about employments, stock market values and sales for firms are obtained from the third source.

9.1.1 Patent Assignment Data (PAD)

The patent assignment data is obtained from the publicly available United States Patent and Trademark Office (USPTO) patent assignment files hosted by Google Patents Beta. These files contain all records of changes made to U.S. patents for the years 1980-2011. The files are parsed and combined to create the dataset. The following variables are kept:

- **Patent number:** The unique patent number assigned to each patent granted by the USPTO.
- **Record date:** Date of creation for the record.
- **Execution date:** Date for the legal execution of the record.
- **Conveyance text:** A text variable describing the reason for the creation of the record. Examples are: "Assignment of assignor's interest", "Security Agreement", "Merger", etc.
- **Assignee:** The name of the entity assigning the patent (i.e., the seller if the patent is being sold).
- **Assignor:** The name of the entity to which the patent is being assigned (i.e., the buyer if the patent is being sold).

- **Patent application date**: Date of application for the patent.
- **Patent grant date**: Date of grant for the patent.
- **Patent technology class**: The technology class assigned to the patent by the USPTO according to its internal classification system.[4]

The entries for which this information is inaccessible are dropped from the sample.

During the parsing process, the following are done:

- Only transfer agreements between companies are kept.
- Only utility patents are kept; entries regarding design patents are dropped.

Following the process described above, and after dropping duplicate entries, the number of records left in the sample is 966,427. Using the string variable that states the reason for the record, all reassignments that are not directly related to sales are dropped (for instance, mergers, license grants, splits, mortgages, court orders, etc.)

In order to create an even more conservative indicator of patent reassignments, a company name-matching algorithm is employed, so that marking internal transfers as reassignments can be avoided, where patents are moved within the same firm, or between the subsidiaries of the firm. The idea behind the company name-matching algorithm is to clean the string variables for the assignor and the assignee of all unnecessary indicators and company type abbreviations. If the cleaned assignor and assignee strings are equal, the type of the record is changed to internal transfer, provided that it was identified as a reassignment before.

The pseudo-code for the algorithm is as follows:

1. All letters are made upper case.
2. The portion of the string after the first comma is deleted. (e.g., AMF INCORPORATED, A CORP OF N.J. becomes AMF INCORPORATED)

---

[4] This variable is not used, however, to represent the technology class for a patent, as is discussed below.

3. If the string starts with "THE", the first 4 characters are deleted.

4. All non-alphanumeric characters are removed.

5. Trailing company identifiers are deleted if found. The string goes through this process 5 times. The company identifiers are the following: B, AG, BV, CENTER, CO, COMPANY, COMPANIES, CORP, CORPORATION, DIV, GMBH, GROUP, INC, INCORPORATED, KG, LC, LIMITED, LIMITEDPARTNERSHIP, LLC, LP, LTD NV, PLC, SA, SARL, SNC, SPA, SRL, TRUST, USA, KABUSHIKI, KAISHA, AKTIENGESELLSCHAFT, AKTIEBOLAG, SE, CORPORATIN, CORPORATON, TRUST, GROUP, GRP, HLDGS, HOLDINGS, COMM, INDS, HLDG, TECH, and GAISHA.

6. If the resulting string has length zero, that string is declared as needing protection. Some examples that are protected by this procedure: "CORPORATION, ORACLE", "KAISHA, ASAHI KAISEI KABUSHIKI", "LIMITED, ZELLWEGER ANALYTICS".

7. The algorithm is re-run from the beginning on the original strings with one difference: The strings that are declared as needing protection skip the second step.

At the end of this process, the number of records that correspond to patent reassignments is 767,815.

9.1.2 USPTO Utility Patents Grant Data (PDP)

The patent grant data comes from the NBER Patent Database Project (PDP), and contains data for all the utility patents granted between the years 1976-2006. How the PDP and PAD are linked to each other is discussed later on.

9.1.3 Compustat North American Fundamentals (Annual)

The Compustat data for publicly traded firms in North America between the years 1974-2006 is retrieved from Wharton Research Data Services. The Compustat database and the

NBER PDP database are connected using the matching procedure provided alongside the PDP data. Extensive information on how the matching is done can be found on the project website.

9.1.4 Connecting PAD and PDP Data

There are two different questions of interest, which require combining the Patent Database Project data with the Patent Assignment Data. The first analysis is on whether a patent is ever reassigned (i.e. sold) over its entire lifetime, and what determines the probability of this event. For this purpose, it is only necessary to connect the information from PAD to the firm which applied for the patent. This is easily done by using the unique patent number each patent is given by USPTO.

The second question involves the change in match quality of the patent when it is traded between two firms. In this case, one needs to know the characteristics of both the assignor and the assignee firm for each reassignment record in the PAD dataset. However, there is no existing connection established between the PAD and PDP datasets. To connect these datasets, the company name-matching algorithm described earlier is employed.

9.2 Variable Construction

9.2.1 Patent-to-Patent Distance Metric

In order to construct a topology on the technology space empirically, it is necessary to create a distance metric between different technology classes. Such a metric enables one to speak about the distance between two patents in the technology space, and leads to the construction of more advanced metrics.

The first 2 digits of the IPC (International Patent Classification) codes of a patent are chosen to indicate its technology class. The IPC code used for a patent is taken from the PDP data and differs from the classification scheme employed in the PAD data. It should be noted that the PDP data actually contain more than a single IPC class for a single patent in some cases, since the IPC codes were assigned using a concordance between the IPC and the internal classification system of USPTO. The IPC code provided in the PDP file with assignees is used in such cases, which is unique for each patent.

As discussed in the main text, a plausible distance metric between patent classes can be generated by looking at how often two different technology classes are cited together. Formally:

$$d(X,Y) \equiv 1 - \frac{\#(X \cap Y)}{\#(X \cup Y)}, \text{ with } 0 \leq d(X,Y) \leq 1.$$

where $\#(X \cap Y)$ denotes the number of patents that cite technology classes X and Y together, and $\#(X \cup Y)$ denotes the number of patents which either cite X or Y or both.

9.2.2 Definition of a Firm in the Data

There are four different entity identifiers in the PDP dataset. The USPTO assignee number is the identifier provided by USPTO itself, but the creators of the PDP have found that it is not very accurate. A single assignee might have many different USPTO assignee numbers. The PDP uses some matching algorithms on the names of the assignees to create a more accurate assignee identifier, called PDPASS. They also link the patent data to Compustat data. Compustat has an identifier called GVKEY. However, these refer to securities rather than firms. So a single firm might be represented by many GVKEY's. For this reason, they use a dynamic matching algorithm again to link all GVKEY's to certain PDPCO's, which is a unique firm identifier that is created by the authors of the project. They create this identifier in order to be able to account for name changes, mergers & acquisitions, etc. This paper follows the same procedure.

9.2.3 Patent-to-Firm Distance Metric

In order to measure how close a patent is to a firm in the technology spectrum, a metric is necessary. However, throughout their lifetimes firms register patents in multiple technology classes. Hence the patent-to-patent distance metric is insufficient for this purpose. One possible way to construct a patent-to-firm distance metric is to compare a patent to the past patent portfolio of the firm. The distance measure between each patent a firm registered in the past, and the new patent in question can be calculated using the patent-to-patent distance metric offered earlier. The distance between the firm and the patent should be a function of these separate distances. Equation (1) defines a parametric family of distance measures indexed by ι. The analysis is conducted for several values of ι.

9.2.4 Creating the Patent Stock Variable for Compustat Firms

As argued in Hall, Jaffe and Trajtenberg (2005), the citation-weighted patent portfolio of a firm is a plausible indicator of the intangible knowledge stock of a firm. The authors demonstrate that this measure has additional explanatory power for the market value of a firm above and beyond the conventional discounted sum of R&D spendings of a firm, since R&D is a stochastic process which can succeed or fail; whereas patents are quantifiable products of this process when it is successful. Furthermore, it is revealed that the number of citations a patent receives is a fine indicator of the patent's worth, increasing the market value of a firm at an increasing rate as the number of citations go higher.

Since all the future citations to a patent cannot be observed at any given date, the citations variable suffers from a truncation problem. There are also technology class and year fixed effects to consider. All of these issues are thoroughly investigated by Hall, Jaffe and Trajtenberg (2005); they provide a variable called *hjtwt* in order to correct the citation number of each patent in the PDP dataset. This study uses their correction method. In the end, a corrected citations number for each patent is obtained. In order to create the patent stock variable of a firm (PDPCO), the corrected citations number across all the patents of a firm are added together at each year. This variable is called *patent stock*.

In addition to the patent stock, the corrected citations number across all the patents of a firm, multiplied by the patent-to-firm distance generated at the date of the patent's inclusion into the portfolio are also added together to create a new variable. This variable quantifies the overall waste of patent stock caused by the mismatch between the technology

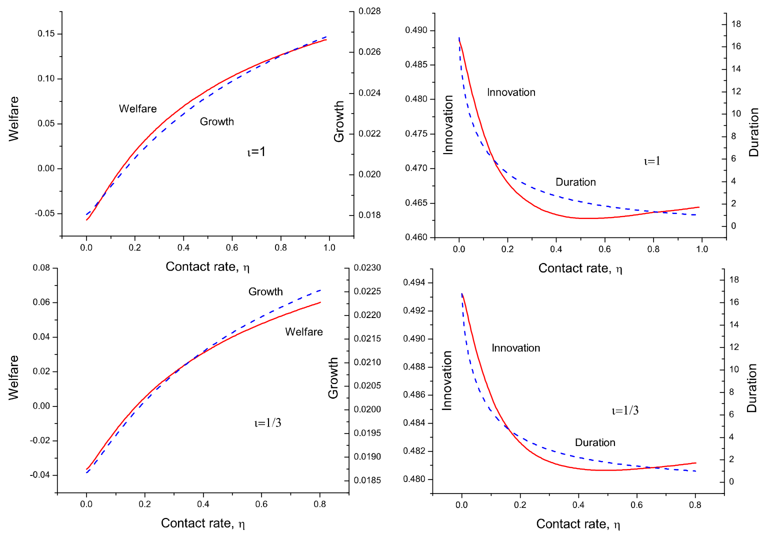

Figure 10: The impact of an increase in contact rate for matching on sales duration, growth, innovation and welfare (for $\iota = 1/3$ and $\iota = 1.0$).

class of the patents and the firm. This variable is expected to have a negative effect on firm moments such as real sales and market value of equity. The variable is called *distance adjusted patent stock*.

9.3 Robustness Analysis

The behavior of the model does not change significantly as ι changes. For example, Figure 10 illustrates how the model behaves as the contact rate for matches, η, changes, for $\iota = 1/3$ and $\iota = 1.0$ (average distance). As can be seen, the results closely resemble those reported in Figure 7 for the baseline model.

Table 9 exhibits the performance of the model in matching the calibration targets. Comparing Table 9 to Table 5 reveals there is no noticable difference in the performance.

Table 9: Calibration Targets - Robustness 1

Target	*U.S. Data*	$\iota = 1$	$\iota = 1/3$
Long-run growth in output	2.00%	2.00%	2.00%
Ratio of R&D expenditure to GDP	2.91%	2.34%	2.50%
Fraction of patents that are sold	15.6%	15.9%	13.3%
Average duration until a sale	5.34 yrs.	5.91 yrs.	4.53
Sale duration, c.v	0.82	1.08	1.10

Likewise, the empirical findings are qualitatively and quantitatively similar when the distance measures created by $\iota = 1/3$ and $\iota = 1.0$ are employed. The significance of the results are maintained, and the coefficients and the ratios of interest are similar in magnitude. Tables 10, 11, and 12 show the regression results with the different metrics.

Table 10: Firm Moments - Robustness

Variable	*log real sales*		*log market value*	
	$\iota = 1$	$\iota = 1/3$	$\iota = 1$	$\iota = 1/3$
log patent stock	0.191***	0.194***	0.037***	0.039***
	(0.008)	(0.008)	(0.008)	(0.008)
log dist-adj pat stock	−0.006***	−0.008***	−0.018***	−0.021***
	(0.003)	(0.003)	(0.003)	(0.003)
log employment	0.936***	0.936***	0.728***	0.728***
	(0.008)	(0.008)	(0.008)	(0.008)
intercept	yes	yes	yes	yes
year	yes	yes	yes	yes
firm fixed effect	yes	yes	yes	yes
Obs.	23,028	23,028	36,094	36,094
R^2	0.96	0.96	0.92	0.92

Standard errors are reported in parentheses.

*10%, **5%, ***1% significance.

TABLE 11: PATENT RESALE DECISION - ROBUSTNESS

Variable	*Indicator (=1 if sold)*	
	$\iota = 1$	$\iota = 1/3$
distance	0.0228***	0.0193***
	(0.001)	(0.001)
patent quality	0.0004***	0.0004***
	(0.000)	(0.000)
log (size of patent portfolio)	−0.0163***	−0.0159***
	(0.000)	(0.000)
intercept	yes	yes
year	yes	yes
firm fixed effect	yes	yes
Obs.	2,564,305	2,564,304
R^2	0.4197	0.4198

Standard errors are reported in parentheses.

*10%, **5%, ***1% significance.

TABLE 12: DISTANCE REDUCTION ON RESALE - ROBUSTNESS

Variable	Change in distance $d(p, f_b) - d(p, f_s)$	
	$\iota = 1$	$\iota = 1/3$
intercept	−0.176***	−0.123**
	(0.056)	(0.067)
year fixed effect	yes	yes
seller fixed effect	yes	yes
Obs.	25,170	25,170
R^2	0.4210	0.3626

Standard errors are reported in parentheses.

*10%, **5%, ***1% significance.

The three different sale duration distributions are created using the following definitions for sale duration:

1. The difference between the sale date and the application date for transactions between 1980-2012.

2. The difference between the sale date and the date of receiving the first citation for transactions between 1980-2012.

3. The difference between the sale date and the grant date for transactions between 1980-2000.

The results of the calibration exercise can be found on Table 13. The model achieves a good fit despite changes in the shape and the mean of the sale distribution. Likewise, the results of the quantitative exercises do not vary by much. For comparison, the welfare contribution of the market for ideas was 4.8% in consumption equivalent terms in the baseline exercise. This number is found to be 4.8%, 4.9%, and 4.5% for the sale duration distributions calculated using the application date, the first citation date, and the grant date with the smaller sample respectively. The other magnitudes change more or less in the same proportion between these welfare numbers, and are qualitatively the same.

Table 13: Calibration Targets - Robustness 2

Target	*Data*	App.	*Data*	First Cit.	*Data*	1980-2000
Long-run growth in output	2.00%	2.00%	2.00%	2.00%	2.00%	2.00%
Ratio of R&D expenditure to GDP	2.91%	2.41%	2.91%	2.41%	2.91%	2.42%
Fraction of patents that are sold	15.6%	15.2%	15.6%	15.4%	15.6%	14.1%
Average duration until a sale (yrs.)	6.06	5.27	5.09	5.21	6.70	5.57
Sale duration, c.v	0.79	1.09	0.80	1.10	0.72	1.09