The Market for Intellectual Property: Evidence from the Transfer of Patents

Carlos J. Serrano*

University of Minnesota
and Federal Reserve Bank of Minneapolis

JOB MARKET PAPER
January, 2005

Abstract

This paper studies the economics of intellectual property transfer, using new data on patent transfers. First, it presents the key stylized facts of the transfer of patents from a dataset I have compiled. The main stylized facts are four: (i) Nearly 20% of all U.S. patents are traded at least once over their life cycle. (ii) Better patents, as defined as those receiving more citations, are more likely to be traded. (iii) The transfer rate, defined as the proportion of traded patents conditional on renewal, monotonically decreases following the date of issue with the exception that it jumps up immediately after a patent is renewed. (iv) Previously traded patents, and in particular those traded recently, are more likely to be retraded and less likely to expire. Second, this paper develops a model with costly technology transfer to interpret the stylized facts. The model has two mechanisms, a selection effect and a horizon effect. The cost of technology transfer creates a selection effect so that better patents are more likely to be traded. This effect accounts for the discontinuous increase in the transfer rate after renewal, and also for why traded patents are more likely to be retraded and less likely to expire. There is a horizon effect that explains why transfer rates decrease as the patent gets closer to its expiration date. A shorter horizon implies less time to amortize the cost of technology transfer. Finally, I structurally estimate the parameters of the model to quantify what are the gains from trade in the market for patents. I find preliminary estimates that show that the existence of this market adds 5.08% to the average value of a patent.

*I am especially grateful to Thomas Holmes and Sam Kortum for their advice and patience. I would also like to thank Chris Laincz, Michele Boldrin, Marcel Boyer, Zvi Eckstein, Nicolas Figueroa, Mike Golosov, Katherine Lande, David Levine, Antonio Merlo, Valeriu Omer, Jennifer O'Reilly, Jim Schmitz, Linda White and participants at the IV Villa Mondragone Workshop in Economic Theory and Econometrics for their comments. I also acknowledge financial support from the Bank of Spain Graduate Fellowship. All errors are mine. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

DURING THE LAST TWO decades, the value of patents, trademarks, copyrights, and other intellectual property assets has surged to become a large part of the wealth portfolio of firms today. In the early 1980's intangible assets represented 38% of the portfolios of U.S. firms. In the mid 1990's this share rose to 70% (WIPO (2003)). Similarly, the transfer of intellectual property assets has become an important source of technology for firms. Nevertheless, the market for intellectual property is not well understood. Previous literature has been hampered by a lack of comprehensive data on how intellectual property assets are traded. This paper is the first systematic work that presents, account for the trading of patents over their life cycle, and estimates the what are the gains from trade in the market for patents. I find that the market for patents adds 5.08% to the intrinsic value of a patent.

My work is distinct to previous literature in that it makes use of data on the transfer of patents. I have derived four key facts using this data. First, a large fraction of patents are traded. Nearly 20% of all U.S. patents issued to small innovators (i.e., firms that were issued no more than 5 patents in a given year) are traded at least once over their life cycle. Second, better patents are more likely to be traded. Logistical analysis shows that an extra citation received by a patent increases the log of the probability of being traded by 0.013 units.[12] Third, the transfer rate varies over the life cycle. It monotonically decreases with the exceptions of the renewal dates and the application period prior to issue of a patent. In particular, immediately after renewal, transfer rates discontinuously increase. Fourth, patents that have been previously traded, and in particular if they have been traded recently, are more likely to be retraded and less likely to expire. These patterns are robust to different patent cohorts and industries.

There is an extensive empirical literature investigating patent data (BLS (1989), Griliches (1979, 1992), Pakes and Griliches (1984), and Hall, Griliches and Hausman (1986), Jaffe, Henderson and Tratjenberg (1993) and Tratjenberg, Henderson and Jaffe (1997), Hall, Jaffe and Tratjenberg (2001), Hall, Jaffe and Tratjenberg (2004) and others). My work is different in that it uses data on transfers of patents. The U.S. patent office registers transfers of patents in the same way that counties register the transfer of houses. As I show here,

[1]Each patent when granted lists references to previous patents, that is citations made. Instead, citations received by a patent is the number of times that this patent has been referenced by other patents. Previous empirical studies on patents have found that citations received by a patent is a good predictor of its value.

[2]The regression I have run consists of a dependent variable that is the decision of whether to trade or not; and independent variables such as a dummy for each patent age and the number of citations received by a given patent age (i.e., ex ante citations received). The coefficient of the variables of ex ante citations received is 0.013. I find that all estimates are statistically significant with a 95% confidence interval.

the market for trade in patents is large. To see this market, the reader is invited to check out "*buypatents.com*". In my study, I make use of all the records of titles transferred and link this information to the basic patent data (e.g., patent's grant date, renewals, citations received, etc.) that others have used.

The dataset I have constructed is a panel with the histories of trades and renewal decisions for patents granted since the early 1980's. In addition, it contains characteristics such as citations received,[3] industry to which the patented technology belongs, size of the firm that was the owner at the grant date, and other relevant information.

There is also extensive work in the theoretical literature on patents. The starting point for my theory is Pakes (1986). He examines the problem of a patent owner deciding in each period whether or not to pay the renewal fee and thereby extend the life of a patent. The contribution of my paper is to introduce into the model, in each each period, an alternative potential owner who may have a greater valuation for the patent than the owner at the beginning of a given period. To transfer a patent to a new owner involves a resource cost, a transaction cost. In summary, whereas Pakes' model has one margin, should the patent owner pay the fee for renewing the patent, my model has a second margin, should the cost of technology transfer be paid to reallocate the patent to an alternative owner.

The intuition for the model is simple. Initially, patents are granted to a fraction of firms. The rest of the firms are potential buyers. At every period, patents are traded because some firms are more productive than others in the use of a given patent. However, any gain from trade in transferring a patent to the potential buyer must be weighed against the resource cost of technology transfer.

In the model, there are two mechanisms. First, the cost of technology transfer creates a selection effect so that better patents are more likely to be traded. This selection effect accounts for the discontinuous increase in the transfer rate after renewal, and also for the reason that traded patents are more likely to be retraded and less likely to expire. Second, there is a horizon effect that explains why transfer rates decrease as patents get closer to their expiration date. A shorter horizon implies less time to amortize the cost of technology transfer.

The parameters of the model are estimated using the simulated general method of moments to fit the proportion of patents that are traded conditional on renewal and the proportion of patent that are allowed to expire conditional on renewal. I find that the average value of patent is $US (2003) 86,782, however patent returns are very skewed. The fixed per patent cost of technology transfer is estimated to be substantial, 29% of the

[3]The size of a firm is defined as the number of patents that were granted to the firm who was the owner of the patent at the time that the patent was granted.

average value of a patent. The existence of the market for patents accounts for 5.08% of the average value of a patent.

The previous literature on markets for intellectual property can be summarized in three groups. One strand aims to demonstrate evidence of the existence of this market. The method generally used has been the analysis of industry case studies, such as the works collected in Arora, Fosfuri and Gambardella (2002). In addition, a sequence of papers by Lamoreaux and Sokoloff (1997, 1999) provide an account of organized markets for technology in the 19^{th} and early 20^{th} century, prior to the growth of in house R&D laboratories by large firms. The second strand of the literature has suggested the existence of potential gains from specialization and diffusion of technology (Arrow (1962), Arora, Fosfuri and Gambardella (2002)). The third strand has analyzed the limitations of the market, such as the appropiation problems in the transfer of knowledge (Arrow (1962a), Teece (1986) and Williamson (1991)) and the cost of transfer of technology (Teece (1977)).

Finally, this work opens new avenues of research. First, to study the sources of innovation and to characterize who are the buyers and sellers of technology. In particular, to trace the flow of technology transfer and to analyze whether small firms specialize in the creation of innovations that eventually are sold to their larger counterparts. Second, to assess the gains from trade derived in the market for patents. Third, to examine to what extent a higher level of patent protection has facilitated specialization and, consequently trade in patents. Lastly, to evaluate the use of taxation on intellectual property transfer to promote innovation. These questions have not been previously empirically addressed due to a lack of data on how patents are traded.

This paper is organized as follows. Section 2 explains the data and presents the stylized facts. Section 3 develops the model. Section 4 solves the model and links the stylized facts with the results of the model. Section 5 present the estimation strategy and discuss identification. Section 6 shows what are the estimation results, in particular it quantifies what are the gains from trade in the market for patents. Section 7 concludes the paper. Finally, data summary tables are in Appendix 1, and all proofs are included in Appendix 2.

2 Data

A patent for an invention is the grant of a property right to the inventor in order to exclude others from making, using, or selling the invention. The life cycle of a patent begins with the application date.[4] The grant date, when the protection period begins, may be up to three or four years later, though the average is 2.5 years. By the end of years 4, 8 and 12 renewal fees of approximately $1000, $2000 and $3000 must be paid, otherwise the patent expires.[5] Such renewal events have been studied for patents granted in European countries in an extensive and important literature (Pakes (1986), Lanjouw (1998) and others).

Another event that can happen in the life cycle of a patent is what the U.S. patent office calls "reassignments", and what I will call a "transfer" or "trade". In principle, the event can happen many times during the life of a single patent. The U.S. patent office maintains a registry of these events. I have obtained these records for all transfers that occurred from 1981 to 2002, of which there were 1,041,083. The records have information about patent numbers, making it possible to merge the patent level data on renewals and citations that has been used in the previous literature. The details of the procedures I used to deal with the transfer data are explained in Serrano (2004).

A particular issue I treat in detail in Serrano (2004) is that some of the transfers recorded with the patent office are administrative events, like a name change, as opposed to a true economic transfer between two distinct parties. Fortunately, for each transaction there is a data field that records the "brief", which is the nature of the trade. I separate out traded patents where the reason is a name change, a security interest, a correction, etc. The remaining accounts for 508,756 patents, 52% of all traded patents.

A second issue is that in cases where there is a merger between two large companies, patents are traded in large blocks. When Burroughs Corporation merged with Sperry Corporation to create Unisys Corporation in September 1986, this event appears in my data as transactions totalling 2261 patents (the largest single transaction includes 1702 patents). My theoretical analysis will focus on decision making at the *patent* level. There are costs and benefits of transferring a particular patent. Obviously, in a wholesale trade such as Burroughs merging with Sperry, the decision making is not at the level of a single patent. To parallel my focus in the theory, in my empirical analysis I focus on *small innovators*. In doing so, the economic forces that I highlight will be more salient than in transactions involving the likes of Burroughs or Sperry.[6] In addition, small innovators are

[4]The term of new patents applied for prior to 1995 was 17 years from their grant date. This term was subsequently modified to 20 years from the date in which the patent application was filed.

[5]The USPTO states that the renewal fee by the end of years 4, 8 and 12 since the grant date of the patent are respectively, US $890, $2,050 and $3,150 as of 2003. The USPTO began charging renewal fees in 1984 on patents applied for after December 12, 1980.

[6]A companion paper, "Measuring the Transfer of Patents" shows that patents granted to large corporations are more likely to be traded for other reasons than the technology that they represent. For instance, they can be recorded as a result of large acquisitions pursued to increase the buyer's market share in a particular product, etc.

interesting in their own right, given the importance they play in the innovation process (Arrow (1983), Acs and Audretsch (1988)). Indeed, I operationalize this focus on small innovators by restricting attention to patents granted to firms with no more than 5 patents granted to them that year.

The dataset I have compiled is a panel of patents detailing their histories of trade and renewal decisions. The panel contains patents that were applied for after December 12, 1980 and issued since January 1, 1983 to U.S. or foreign businesses.[7] In addition, it has characteristics such as citations received, industry to which the patented technology belongs, size of the firm that was the owner at the grant date, and other relevant information.

The panel includes 491,611 patents granted to small innovators. This sample contains a third of all granted patents to U.S. or foreign businesses.

The next section presents the key stylized facts of the transfer of patents.

2.1 Stylized Facts

This section presents the basic facts that describe the underlying quality of traded patents, how the transfer rate varies over the life cycle of a patent, and the effects of a transfer on the renewal and trading decision.[8] The key facts are the following:

1. A large fraction of patents is traded.

2. Patents with higher quality are more likely to be traded.

3. The transfer rate[9] monotonically decreases following the first year after a patent has been granted, with the exceptions of the renewal dates. Immediately after renewal, the transfer rate discontinuously increases. Moreover, the transfer rate increases during the application period of a patent.

4. Patents that have been previously traded, in particular those recently traded, are more likely to be retraded and less likely to expire.

[7]Patents applied for after December 12, 1980 are subject to renewal fees. Patents tend to be granted two years after their application date (I only have data on application of patents that were eventually granted). To create a comphensive sample I consider January 1, 1983 as the starting grant date of the patents contained in the panel. Finally, issued to U.S. or foreign business means that at the date the patent was granted, the owner was a U.S. or foreign business.

[8]For explanation, the age of a patent is defined as follows. Its age when it is traded is the number of years between the trade date and the grant date. In particular, if a patent was traded during its second year of life (e.g., 17 or 22 months since being issued), I consider that the patent was traded at age 2.

[9]Transfer rate at a given patent age is defined as the proportion of patents that are traded conditional on having survived up to that period.

Fact 1: A large number of patents are traded, and these patents are found to be of higher quality when measured by ex post citations received.[10] Table 1 shows that nealy 20% of the patents issued to small innovators are traded at least once over their life cycle.[11] In addition, these patents represent about 35% of all ex post patent citations (i.e., weighted by ex post citations received).[12]

Table 1: Percent of Patents that are Traded

Unweighted	19.8
Weighted	34.3

Fact 2: Patents with a higher number of ex ante citations received are more likely to be traded. Logistical analysis shows that an extra citation received by a patent increases the log of the probability of being traded by 0.013 units. [13]In particular, patents with 1 ex ante citations received at age 8 have an estimated probability of being traded of 0.0177 at age 8, this probability jumps to 0.038 if ex ante citations received are 60, and it spikes to 0.063 if ex ante citations received are 100. The following table summarizes the estimated transfer rate conditional on the number of ex ante citations received by a given patent age.

Table 2: Transfer Rate Conditional on Citations Received

Age	Unconditional	Ex Ante Citations Received					
		0	1	10	20	60	100
1	0.0261	0.0260	0.0264	0.0297	0.0337	0.0561	0.0919
8	0.0188	0.0175	0.0177	0.0199	0.0227	0.0380	0.0630
17	0.0085	0.0073	0.0074	0.0084	0.0095	0.0162	0.0272

In addition, better patents are less likely to expire. Evidence from citations received over a patent's life cycle shows that patents with higher ex ante citations received are less likely to be allowed to expire. This feature confirms the robustness of using dynamic citations as a measure of the quality of patents. Logistical analysis predicts that an extra citation received decreases the log of the probability of being expired by 0.045 units. The following table presents the estimated expiration rate conditional on the number of ex ante citations received by each renewal date.

[10]Ex post citations received are defined as the total number of citations received measured by the end of the last year that the patent is present in the panel. Instad, ex ante citations received by a given patent's age are the cumulative number of citations received since the patent's grant date until the given patent age.

[11]Furthermore, actual sales of patents is a lower bound on transactions in intellectual property since no broad data on licensing is available.

[12]Previous literature have found ex post citations received as a good proxy of the value of a patent.

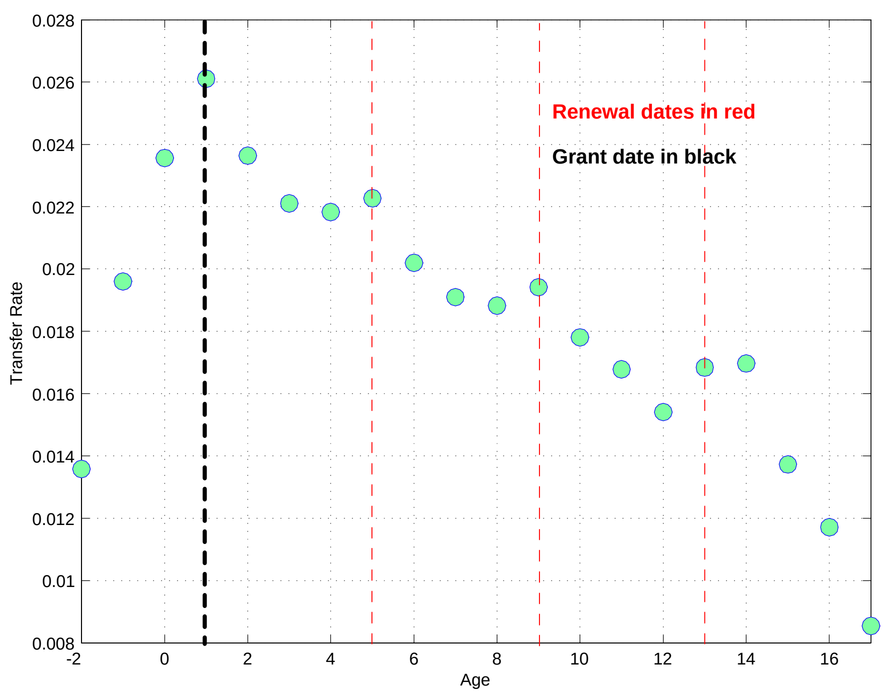
[13]Details about the logistical analysis can be found in a companion paper: "Measuring the Transfer of Patents," manuscript, at http://www.econ.umn.edu/~carles/research.htm

Table 3: Expiration Rate Conditional on Citations Received

Age	Unconditional	Ex Ante Citations Received					
		0	1	10	20	60	100
5	0.1809	0.1990	0.1920	0.1369	0.0920	0.0165	0.003
9	0.2833	0.3342	0.3243	0.2427	0.1698	0.0329	0.006
13	0.3202	0.3981	0.3874	0.2969	0.2123	0.0429	0.007

Fact 3: The transfer rate is hump-shaped as a function of the age of a patent (see Figure 1). This rate monotonically decreases from the grant date until the end of a patent's life except at the renewal dates. In these periods, the rate increases abruptly. For instance, the proportion of traded patents drops from 2.7 to 2.2% respectively, from age 1 to age 4. Then in age 5 it increases to 2.3 and in age 6 drops again to 2.05%. This evidence is consistent for all three renewal dates in which renewal fees are due according to the U.S. patent system. Before a patent is granted, that is in the application period, which lasts on average 2.5 years,[14] the transfer rate increases.

Figure 1: Transfer Rate Conditional on Renewal



Fact 4: This fact focuses on the effects of the trading decision on the retrading and the future renewal decision of the patent. Table 3 shows a combination of results that explain

[14] In addition, 90% of all issued patents contained in the sample are granted by their third year after the application date.

how the trading decision and its timing show the difference between previously traded and non-traded patents.

In particular, columns labeled as "Not Traded" and "Traded" of table 4 show that among patents of the same age, traded patents are more likely to be retraded and less likely to expire. We see that previously traded patents are twice more likely to be retraded. With respect to the renewal decision, traded patents are about 5 percentage points less likely to expire at each of the renewal dates.

In addition, columns labeled as "Traded 1y ago" and "Traded 4y ago" show that these differences are even more striking when we consider the timing at which a previous trade took place. For instance, patents traded a year before a renewal date are about half as likely to expire at that renewal date than patents traded four years ago. Finally, patents traded one year ago are twice as likely to be traded than patents traded four years ago.

Table 4: Proportion of Patents Traded and Expired

	Age	Not Traded	Traded	Traded 1y ago	Traded 4y ago
Expired	5	0.186	0.138	0.071	0.150
	9	0.293	0.240	0.118	0.229
	13	0.330	0.288	0.156	0.275
Traded	1	0.024	0.052	0.052	-
	4	0.018	0.042	0.050	0.037
	8	0.013	0.036	0.048	0.038
	12	0.010	0.025	0.033	0.026

The next section develops a model that interprets the key stylized facts.

3 A Model of Patent Trades

The main objectives of the model are to shed light on the underlying quality of patents that are traded, and to study how the cost of technology transfer determines the probability of being traded conditional on renewal as a function of the age of the patent (i.e., the transfer rate as a function of patent age).

To do so, we consider an economy with time indexed by $a = 1, ..., L$ where $L < \infty$ represents the finite life of a patent.[15] The economy is populated by a large number of firms $\mathcal{F} < \infty$. In period 1, the economy is also endowed with a large number of patents

[15] Notice that periods and the age of a patent are interchangeable.

$K < F$, which are randomly paired to firms. Some firms are more productive and can obtain more revenues than other firms from a given patent. Thus at the beginning of every period a fraction of all the firms holds exactly one patent, and the rest of the firms can be alternative owners by making succesful acquisition offers.

Firms obtain their revenue from the patents they own and each patent is owned by at most one firm.[16] For simplicity, I also assume that each firm holds at most one patent.

Firms maximize profits. Profits are defined as the expected discounted value of a sequence of per period patent returns x_a minus a renewal fee c_a. The discount factor is β. The payment of the fee extends patent protection. If the fee is not paid, then returns are zero thereafter. Consequently, the value of a firm is exactly the value of its patent. Therefore, the growth of a firm is the growth of patent returns.

The timing is as follows. At the beginning of every period, firms know the per period return of their patent in case it is allowed to expire ($zero$), kept (x), or sold (y) to a potential buyer. Firms choose whether to sell, keep or let the patent expire. Next, the per period return of the patent, let's us $z \in \{0, x, y\}$ be that return, is collected by the new or old owner if the patent has been sold or kept. Finally, at the end of the period, the returns for next period become known, the owner of the patent meets a new potential buyer, and an acquisition offer is received.

In the model, there are two sources that account for how per period returns evolve over the life cycle of a patent. Patent returns can grow within the firm and between firms. First, the internal process of returns takes place within the firm. This process represents either the arrival of new potentially applications about the innovation than increase its value, or the depreciation of returns due to the discovery by other firms of similar technologies. Second, the external growth of returns occurs if the patent is acquired by another firm. This accounts for the possibility that some firms are more efficient and can obtain more revenues from the same patent.

Growth of Returns Within the Firm The internal growth of returns is modeled as a stochastic process that captures all sources that affect the growth of patent returns other than the ones associated with the efficiency gains attained from the transfer of patents. For instance, possible interpretations can be (i) lower growth of returns due to imitation, arrival of superior technologies to produce a similar good; or (ii) higher growth of returns due to the arrival of new applications that enhance the returns of the innovation, or learning about the product market of the innovation.

[16] Licensing of patents is in the background. Licensing affects the per period revenue of a patent but not its ownership.

These characteristics motivate a process of returns within the firm that might depend on the age of the patent, a, and the per period patent return, z_a. Notice that z_a is determined right after the decision of selling, keeping or renewing the patent. Thus, z_a can be either zero, x_a, or y_a if the renewal fee was not paid (i.e., patent is allowed to expire), the patent was kept, or the patent was traded. The next period returns are

$$x_{a+1} = g_a^i z_a \qquad a \in \{1, ..., L-1\}$$

where $g_a^i \in [0, B^i]$ is a random variable that represents the growth of internal returns. The random variable is distributed with

$$F_{g_a^i}(u^i; z_a, a) = \Pr[g_a^i \leq u^i; z_a, a]$$

where $F_{g_a^i}(.)$ is the distribution function of g_a^i.

For simplicity, I focus on a process in which the growth of returns within the firm is independent of their level. This case helps to disentangle and highlight two important effects present in the data: the selection and horizon effects. In section 4.4, implications of a process of growth of returns that allows for dependence on the level of returns are also studied.

The case for which the growth of returns is independent of their level has a counterpart in empirical industrial organization. Many studies have persistently found evidence for which the growth of firms is independent of their size, known as Gibrat's law. This general case has not been previously considered in the patent literature.[17] However, I argue that in the theory developed in this paper, the growth of a firm is equivalent to the growth of patent returns. Since each firm holds a patent, and patent protection is their unique source of revenue, success, survival and exit for innovative firms and patents have effectively the same essence.

The following assumption specifies Gibrat's law within the model.

Assumption G: The process of returns of a patent follow Gibrat's law if the internal growth of the returns is independent of their level

$$F_{g_a^i}(u^i; z_a, a) = \Pr[g_a^i \leq u^i; a]$$

Growth of Patent Returns Between Firms In this environment some firms are more efficient and can obtain more revenue than other firms from a given patent. Thus, growth

[17] However, linear depreciation, which is a particular case of Gibrat's law, has been previously used in the literature on estimating the value of a patent (see Lanjouw, Pakes and Putnam (1998))

of patent returns between firms is a result of efficiency gains attained through the trading of patents.

In the model, at the end of every period, each holder of a patent meets with a potential buyer. The potential buyer draws an efficiency factor g^e. This efficiency gain represents the efficiency of the potential buyer relating to that of the current owner of the patent. So, at the beginning of every period, the patent return of the potential buyer is defined as

$$y_{a+1} = g^e x_{a+1} \qquad a \in \{1, ..., L-1\}$$

where the random variable g^e is drawn independently and identically from a distribution F_{g^e} with support $[0, B^e]$. Thus, the probability that an efficiency factor is lower than any given number u^e is

$$F_{g^e}(u^e) = \Pr[g^e \leq u^e]$$

The distribution function F_{g^e} is common to all patents at all ages.

However, the diffusion of innovations across the boundaries of a firm is not cost free. An investment must be pursued to make efficient use of the acquired knowledge. This is assumed a fixed cost of technology transfer, τ, independent of the age of the patent and the potential gains from trade. The existence of significant costs of technology transfer is a well known fact documented in the literature of intellectual property transfer. Teece (1977) is an early reference.

Finally, during the meeting between the potential buyer and the owner of the patent, an offer is made. This offer to buy a patent can be summarized by the age of a patent, a, and the patent return of the potential buyer, y_a. Further, any efficiency gain from transferring the patent must be weighted against the cost of technology transfer. For simplicity, it is assumed that the seller gets all the surplus.

The next section studies the problem that a firm solves.

3.1 The Maximization Problem of a Firm

Consider the problem of a firm that holds a patent prior to its a^{th} renewal. At the beginning of any period the holder of a patent knows its current return if respectively, the patent is kept x, it is sold y, or allowed to expire, zero. Consequently, the decision is made and the patent return is $z \in \{x, y, 0\}$.

Define $V(a, z)$ as the discounted expected value of a patent of return z at age a when the firm is commited to pay the renewal fee.

$$V(a, z) = z - c_a + \beta E[\widetilde{V}(a+1, x', y')|a, z]$$

This value is equal to the current return of the patent, z_a, minus the renewal fees, c_a, plus its discounted option value. The next period internal returns are $x' = g_a^i z$, and the external returns are $y' = g^e x'$. In addition, the operator $E[.]$ denotes an expectation conditional on the return of the current owner and the age of the patent. The option value of a patent is defined as

$$E[\widetilde{V}(a+1, x', y') \mid a, z] = \int\int \widetilde{V}(a+1, u^i z, u^e u^i z) \mathrm{d}F_{g_a^i}(u^i; a, z) \mathrm{d}F_{g^e}(u^e)$$

The value of keeping a patent is $\widetilde{V}^K(a, x, y)$. This only depends on the age of the patent and the per period return of the current owner.

$$\widetilde{V}^K(a, x, y) = V(a, x)$$

Instead, the value of selling is $\widetilde{V}^S(a, x, y)$. This is equal to the value of the patent for a firm with current returns y minus the cost of technology transfer τ. For simplicity, it is assumed that the owner of a patent makes a take it or leave it offer to the buyer.[18] Then, the value of selling, which coincides with the price of the patent, is

$$\widetilde{V}^S(a, x, y) = V(a, y) - \tau$$

Finally, the holder of a patent decides whether to sell, keep or let its patent expire by solving the following

$$\widetilde{V}(a, x, y) = \max\{\widetilde{V}^S(a, x, y), \widetilde{V}^K(a, x, y), 0\} \qquad a = 1, ..., L$$

where the letters S, K denote the value of sell and keep and zero is the return if a patent is let to expire.

Prior to the description of the equilibrium of the model, several assumptions are needed to characterize basic properties of the value function. A sufficient condition is that returns are bounded, or that they depreciate with age.

Basic continuity conditions for the process of growth of returns within the firm are assumed (i.e., generality conditions to guarantee the continuity and existence of the value

[18] Alternative bargaining methods do not affect the qualitative results of the model.

function): First, there exists an ε such that $E[{x_a}^{1+\varepsilon} \mid a = 1] < \infty$. Second, $F_{g_a^i}(u^i \; ; z_a, a)$ is continuous in z at every u^i except, possibly, at values of u^i at which $F_{g_a^i}(u^i \; ; z_a, a)$ has a discontinuity in u^i.

In the model it is also considered that the expected growth of returns within the firm decreases with the age of the patent. This is equivalent to assuming that $F_{g_a^i}(u^i; z_a, a)$ is weakly increasing in a. We focus on processes of growth of patent returns such that patents with high returns today are more likely to have high returns tomorrow. This translates to assuming that the $\Pr(zu_a^i \leq \overline{u}|z)$ is weakly decreasing in z, that is first order stochastic dominance in the evolution of per period returns within the firm.

The above conditions will be carried out through the paper (further assumptions are used later in the paper, and they are stated in those cases).

The following Lemma states that the value function of a patent is continuous, weakly increasing in the returns of the patent and weakly decreasing in patent age.

Lemma 1 *The value function $\widetilde{V}(a, x, y)$ is continuous and weakly increasing in the current return of the holder of the patent, x, and the return of the potential buyer, y. The option value $E\widetilde{V}(a+1, x', y'|z, a)$ is weakly decreasing in a.*

Proof. See Appendix. ■

The next section develops the main implications of the model.

4 The Selection and Horizon Effect

This section is devoted to the selection and horizon effects and their role in explaining the stylized facts observed in the data. To do so, the characteristics of the policy functions of the problem of the firm are analyzed. First, I focus on the selection effect. In order to do this, the comparative statics for the case of a fixed patent age are studied. Subsequently, I study the foundations of the horizon effect. This conveys the analysis of how the two cutoff rules vary over the life cycle of a patent. Finally, I examine the case in which the growth of returns within the firm depend on their level.

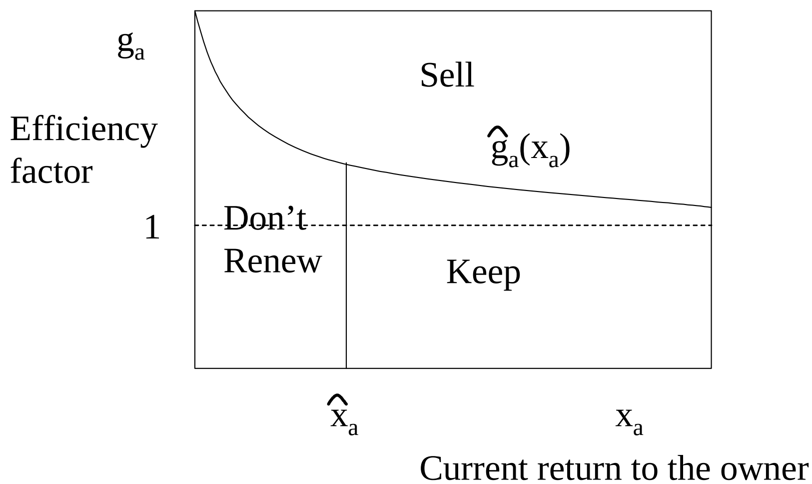
The policy functions are two cutoff rules $\{\widehat{x}_a(\tau)\}_{a=1}^L$ and $\{\widehat{g}_a(x, \tau)\}_{a=1}^L$.[19] They depend on the current return of a patent, x, the return of a potential buyer, y, and the parameters of the model such as the renewal fees c_a, and the cost of technology transfer τ. The policy space can be summarized by these policy rules. First, $\widehat{x}_a(\tau)$ is defined as the patent return

[19] Both cutoff rules also depend on the states (x, y, a), the renewal fees, c_a, and the cost of technology transfer, τ. However, c_a and y have been omitted.

that makes the holder of a patent indifferent between keeping or letting a patent of age a expire. Second, the cutoff $\widehat{g}_a(x, \tau)$ represents the potential external growth in returns that makes a firm indifferent on whether to trade a patent or not.

The two rules describe three regions, E, K, and S in the policy space of patent of age a. These regions correspond respectively to: let the patent expire, keep it or sell it. The regions E and K are separated by a straight line $\widehat{x}_a(\tau)$, which is independent of the external growth of returns g^e. For low current patent returns, that is $x < \widehat{x}_a(\tau)$, the firm chooses between not renewing and selling the patent. So, the cutoff $\widehat{g}_a(x, \tau)$ separates the areas E and S. Finally, for sufficiently high returns, that is $x > \widehat{x}_a(\tau)$, letting the patent expire is not an optimal choice, thus $\widehat{g}_a(x, \tau)$ delineates the K and S regions. The following figure shows a particular example of the policy space, under Gibrat's law.

Figure 2: Policy Space



The next section studies the selection effect. This coincides with the examination of the problem of selling, keeping or letting expire a patent of age a. In other words, it characterizes the underlying properties of the patents that are traded.

4.1 The Selection Effect: Policy Functions for Fixed Age

This section explores the individual heterogeneity in patent characteristics that makes some patents more likely to be traded. In the model, patents are defined by a pair (a, x), where a is the patent age and x is the per period return. The analysis of this section is focused on how ex ante patent returns determine the patent's likelihood of being traded, and similarly its likelihood of being allowed to expire.

4.1.1 Which Patents Are Traded?

Evidence presented in Section 2 shows us a few things. First, patents with a higher number of ex ante citations received are more likely to be traded. Second, traded patents are more likely to be retraded and less likely to expire. These patterns correspond, in particular, to stylized facts 2 and 4 mentioned earlier.

The model can account for these patterns. The mechanism is simple. The cost of technology transfer creates a selection effect so that patents with higher ex ante per period returns are more likely to be traded. In addition, since new owners are more efficient in using the patent, its return increases even further after being traded. Thus, patents of higher quality are more likely to be traded, and traded patents are more likely to be retraded and less likely to expire.

The selection effect arises in the following way. In the model, any patent owner will sell a patent if it receives an offer from another firm with relative efficiency higher than the efficiency that offsets the cost of technology transfer, that is $g^e > \widehat{g}_a(x,\tau)$. So, the probability that a patent is traded concides with the likelihood of receiving such offers, which is $\Pr(g^e \geq \widehat{g}_a(x,\tau))$. This probability, which is a function of age and patent return, is defined as the transfer rate of a patent of age a.

The characteristics of the transfer rate are ultimately determined by those of the cutoff $\widehat{g}_a(x,\tau)$. In other words, if $\widehat{g}_a(x,\tau)$ is decreasing in x, then the probability of being traded increases in x. Therefore, better patents are more likely to be traded if $\widehat{g}_a(x,\tau)$ is decreasing with x.

In order to show that $\widehat{g}_a(x,\tau)$ is decreasing in x, we must consider two separate cases: (i) $x < \widehat{x}_a(\tau)$, and (ii) $x > \widehat{x}_a(\tau)$. This separation is necessary because the function $\widehat{g}_a(x,\tau)$ is defined differently in these two parts of the policy space. In the first case, when patent returns are sufficiently low, that is $x < \widehat{x}_a(\tau)$, the proof is obvious. However, I briefly explain it because it provides a clear intuition of the selection effect.

When per period patent returns are low, the choices of the firm are just whether to sell or let the patent expire. A seller is indifferent to selling or allowing the patent to expire if $\widetilde{V}^S(.) = 0$. We can show that there exist $\widehat{g}_a(x,\tau)$ such that the firm is indifferent between these two choices, i.e., $V(a, \widehat{g}_a(x,\tau)x) - \tau = 0$. The left hand side of the equation is the value of selling a patent, and the right hand side is the value of an expired patent. If x increases, then it must be the case that $\widehat{g}_a(x,\tau)$ decreases with x to keep the equality holding. So, the higher x is, the lower $\widehat{g}_a(x,\tau)$ is to cover the cost of technology transfer. The result we want to show follows because the value function of a patent is weakly increasing in the level of current returns as shown in Lemma 1.

Instead, if per period returns are sufficiently large, that is $x > \widehat{x}_a(\tau)$, then the firm acts at the margin between selling and keeping the patent. The proof is a bit more elaborate than the argument of low returns, but the result still holds. It can be shown that if the internal growth of returns is independent of the level, that is Gibrat's law, then the cutoff $\widehat{g}_a(x,\tau)$ is indeed monotonically decreasing in x. This result means that owners of patents with larger returns demand less external growth of returns in acquisition offers for their patents to be traded. The following proposition summarizes the result.

Proposition 2 *If assumption G holds and $\tau > 0$, then (i) the function $\widehat{g}_a(x,\tau)$ is weakly decreasing for all x, (ii) $\widehat{g}_a(x,\tau) > 1$, (iii) the probability of being traded is increasing with x.*

Proof. See Appendix. ■

The intuition of the proof is not difficult. For instance, consider the case of a myopic firm, which is when the discount factor β is zero. Later in this section the case of positive discount factor will be considered. For a myopic firm, the value of a patent at a given age is just the current returns. Consequently, the value of keeping and selling is the current return x and the return of a potential buyer minus the cost of technology transfer $y - \tau$. A firm is indifferent between both choices if there exists an efficiency gain $\widehat{g}_a(x,\tau)$ such that $y = x\widehat{g}_a(x,\tau)$ holds. Now, fix y. The larger x is, the lower $\widehat{g}_a(x,\tau)$ must be to cover the cost of technology transfer and for the equality to hold. Therefore, patents that are traded are ex ante of higher quality that the average patent.

For the general case in which the discount factor is positive, the argument for a proof is as follows. We must show that for a given efficiency gain between a potential buyer and seller, let's say $\overline{g}_a^e$, and patent age a, the difference between the option value of selling and keeping is weakly increasing in the current return of the patent, z_a. In other words, let us increase the current return of the patent z_a a bit, now we can show that the difference between selling and keeping also increases. Then, there exists a lower efficiency gain g_a^e lower than $\overline{g}_a^e$ such that the difference in expectations (see below) goes back to the level before increasing z_a. Therefore, in order to show that $\widehat{g}_a(x,\tau)$ is decreasing in x, it is sufficient to show that the following expression increases as a function of z_a.

$$E\left[\widetilde{V}(a+1,x',y')|a,g_a^e z_a\right] - E\left[\widetilde{V}(a+1,x',y')|a,z_a\right]$$

However, the proof is a bit subtle. This is due to the fact that, because both option values are weakly increasing in z_a, then the change in the difference with respect to z_a can be ambiguous. In fact, the result depends on the characteristics of the process of internal growth of returns, which is $F_{g_a^i}(u^i; z_a, a)$. Proposition 1 shows that proportional growth of returns, that is Gibrat's law, guarantees that the above difference in expectations is weakly increasing in the current return of the patent. Section 5 explores other processes of internal growth of returns that might depend on the level of returns.

This section has shown that patents with higher ex ante returns are more likely to be traded, and its implications are traded patents are more likely to be retraded and less likely to expire. This result had relied on two key assumption of the model: (i) a fixed cost of technology transfer, and (ii) that gains from trade are relative to the current return of the patent (i.e., the Gibrat's assumption, that is a proportional growth rate, is a particular case of this).

First, if the cost of technology transfer, instead of a fixed cost, is a proportion of x (i.e., $\tau = \alpha x$, where $\alpha \in (0, 1]$), then the result of proposition 1 still holds from any $\beta > 0$. However, if the cost of technology is either zero or proportional to the value of the buyer prior to taxes, $V(a, y)$, then the probability of being traded is independent of the return of the patent (i.e., it is flat with respect to x). So, the implication of these two hypothetical cases would not coincide with the facts observed in the data.

Second, if the process of external growth of returns that allows for trades in patents was in levels rather than in growth rates (i.e., the random draw could be a direct return y, rather than a g^e), then the probability of being traded would be decreasing in the return of the patent. This would contradict the facts observed in the data. Thus, the model shows that arrival offers must be linked to the current return level of a patent in order to explain the data.

Therefore, the model presents minimal assumptions that account for the stylized facts in the data. In particular, it predicts the necessity to incorporate costs of technology transfer. Finally, it shows the need of modeling potential gains from trade conditional on the current return of the patent to explain the stylized facts (i.e., a relative efficiency gain).

4.1.2 What Patents Are Allowed to Expire?

Patents expire because their returns are sufficiently low and not enough to cover the renewal fees. This is an obvious implication of the construction of patent renewal models (see Pakes (1986)).

In the model, the result holds because the option value of a patent in increasing in its return, which was shown in Lemma 1. Any patent of age a , if current returns are $x < \widehat{x}_a(\tau)$ and a potential buyer offer is characterized by an efficiency gain $g^e < \widehat{g}_a(x, \tau)$, then the patent will be allowed to expire.

The next section focuses on the effects of age in the trading and renewal decision of a patent.

4.2 The Horizon Effect: How do Policy Functions Change with Age?

This section studies how the transfer rate of patents varies over their life cycle. In particular, it analyzes the comparatives statistics, with respect to the age of a patent, of the policy functions for a fixed patent return. The interplay between the cutoff $\widehat{g}_a(x,\tau)$ and the age of a patent determines the probability of being traded conditional on survival. This probability might be of interest because it explains the expected efficiency gain from a patent transfer. In addition, the section also looks at how the likelihood to expire changes over the patent's life cycle. This is interpreted by how the cutoff $\widehat{x}_a(\tau)$ affects the probability of being allow to expire as a function of age.

4.2.1 The Transfer Rate of a Patent Over Its Life Cycle

The stylized fact 3 shows that the transfer rate of patents monotonically decreases since their issue date with the exception of the renewal dates. Immediately after renewal, the transfer rate discontinuously increases.

In the model, there is a horizon effect that explains that transfer rates decrease as the patent gets closer to its expiration date. A shorter horizon implies less time to amortize the fixed cost of technology transfer.

The transfer rate of a patent is the probability of being traded conditional on survival. Patents are traded if efficiency gains from purchasing offers are such that $g^e > \widehat{g}_a(x,\tau)$. So the probability of being traded is just $\Pr(g^e > \widehat{g}_a(x,\tau))$, which, in particular, depends on a. In order to show that the transfer rate is decreasing with age, it suffices to prove that the cutoff $\widehat{g}_a(x,\tau)$ is increasing in a for a fixed x.

The argument to show the result can be divided into two parts. First, it is easy to show that the result holds if returns are low, $x < \widehat{x}_a(\tau)$. When returns are low, $x < \widehat{x}_a(\tau)$, the firm's optimal choices are between keeping the patent or allowing it to expire. A firm is indifferent between these two choices if $V(a, x\widehat{g}_a(x,\tau)) = 0$. By Lemma 1, we know that the value of a patent is decreasing in age and increasing in returns. So, if a increases, then it must be the case that $\widehat{g}_a(x,\tau)$ also increases to keep the equality holding. Thus for a fixed low patent return x, offers with higher efficiency gains are required for a trade to take place as a patent gets older.

Second, proposition 2 generalizes the result for all returns and for any period other than the renewal dates. It is assumed that the growth of internal returns is independent of the level.

Proposition 3 *If assumption G holds and $\tau > 0$, then for all x and τ, (i) $\widehat{g}_a(x, \tau)$ is increasing in a, and (ii) the probability of being traded conditional on survival is weakly decreasing in age, a.*

Proof. See Appendix. ■

The argument of the proof when returns are high (i.e., $x > \widehat{x}_a(\tau)$) is to show that, for a fixed efficiency gain g^e, the difference between the value of selling and keeping a patent monotonically decreases, and also converges to zero at most at period L. If the internal growth of returns is independent of the level, then the result follows because both the value functions of keep and sell decrease in average proportionally to their current return level.

4.2.2 The Expiration Rate of a Patent Over Its Life Cycle

This section studies the probability of being expired as a function of the age of the patent. The foundations of this result were shown in Pakes (1986).

Evidence from U.S. patents presented in stylized fact 2 shows that the proportion of patents allowed to expire increase with the age of a patent. In the model, to explain this result it is sufficient to show that the cutoff $\widehat{x}_a(\tau)$ is increasing in the age of the patent. The following proposition shows that.

Proposition 4 *If the renewal fees schedule is weakly increasing, then the cutoff $\widehat{x}_a(\tau)$ is weakly increasing in age. Therefore, the probability of being allowed to expire weakly increases as a function of patent age.*

Proof. See Appendix. ■

More interesting, however, are the effects of the renewal decision on the distribution of patent returns as a function of patent age. This is particularly relevant when the number of renewal dates is low, as it is in the U.S. patent system. In this case, the number of patents that might expire in each of the renewal dates can be substantially large. Thus, immediately after a large fraction of patents expires, the average patent return increases. The next section develops and explores this result in-depth.

4.3 A Horse Race: The Selection Compared to the Horizon Effect

Renewal dates and their implications into the trading decision are interesting events on which to focus.

These events link the results of the model with the data. The model predicts that immediately after renewal, the selection and horizon effects display opposite trends towards the probability for a patent of being traded. On the one hand, right after the renewal date, the average patent return increases, so patents are more likely to be traded according to the selection effect. On the other hand, the horizon effect implies that as the age of a patent increases, patents are less likely to be traded . Therefore, there exists a horse race between the two effects. The race determines the observed proportion of patents that are being traded.

In addition, the model can separately identify the impact of the two effects. To do so, let us consider the possibility of shrinking a period in the model to just less than seconds (i.e., for understanding, a continuous-time model would be a good approximation). If at equilibrium a positive measure of patents expire, then immediately after renewal the distribution of returns dominates stochastically the one existing before the renewal decision. Thus, the selection effect implies that the probability of being traded increases. However, as the time passes by since the renewal date, the horizon effect becomes stronger so that eventually it offsets the selection effect. Therefore, the model helps to disentangle the forces under the selection and horizon effect.

Evidence from the data seen in stylized fact 3 shows that the probability of being traded increases in all renewal dates. Then this probability monotonically decreases until the next renewal date. Therefore, according to the model this is a sign that shortly after renewal the selection effect is sufficiently stronger than the horizon effect, but also the selection effect vanishes fast.

4.4 Level Dependence in the Growth of Returns within the Firm

Previous sections of the paper have focused on the case for which the growth of patent returns within the firm was independent of the level.

However, the existing patent literature that has estimated the value of patents, such as Lanjouw and Pakes, have considered variations of a process of growth of internal returns that also depends on their level. The following assumption defines an explicit stochastic process of growth of returns within the firm with level dependence.

Assumption L: The random variable g_a^i is distributed with a function that depends on the return z_a and patent age a.

$$F_{g_a^i}(u^i; z_a, a) = \left\{ \begin{array}{lll} 0 & \text{if} & u^i < \delta \\ \Pr[g_a^i \leq u^i; z_a, a] & \text{if} & u^i \geq \delta \end{array} \right\}$$

such that the function $F_{g_a^i}(u^i; z_a, a)$ is increasing in z.

If $F_{g_a^i}(u^i; z_a, a)$ does not depend on age, then the process is defined as constant learning (LC). If $F_{g_a^i}(u^i; z_a, a)$ is increasing in a, then the process is defined as diminishing learning (LD). An extreme case of LD, let is call it LDE, is one in which there exist an $\overline{a} < L$ such that $\Pr[g_a^i = \delta; z_a, \overline{a}] = 1$.

This assumption considers a process such that in every period returns either depreciate at a rate $\delta \in (0, 1)$ or grow at a rate $g^i > \delta$.

The term learning has been previously used in the literature on estimating the value of a patent. Learning allows for the possibility of new opportunities or applications that enhance the returns of a patent to be discovered (i.e., so that their growth rate can be higher than depreciation).

The following proposition proves that, if the internal growth of returns also depends on the level as in the learning process specified in assumption L, then $\widehat{g}_a(x, \tau)$ is not necessarily increasing for all a. In particular, it states sufficient conditions so that $\widehat{g}_a(x, \tau)$ is weakly decreasing for sufficiently young patents. The main implication is that the probability of being traded is hump-shaped as a function of the age of the patent.

Lemma 5 *If the support of the random variables g_a^i and g^e is bounded above, then* $\left| \frac{\Delta V(a+1, x', y' | z)}{\Delta a} \right| \leq Q$.

Proof. See Appendix. ■

Proposition 6 *Let $\overline{a}$ be the patent age at which learning completely vanishes. If assumption LDE, the support of the random variables g_a^i and g^e is bounded above, and the $\Pr(g_a^i \leq u^i; z_a, a)$ is sufficiently concave with a, then there exists an age $a^* < \overline{a}$ such that for $a < a^*$ the function $\widehat{g}_a(x, \tau)$ is weakly decreasing in a. In addition, if $a \geq a^*$, then $\widehat{g}_a(x, \tau)$ is weakly increasing in a. Therefore, the probability of being traded conditional on survival is hump-shaped as a function of the age of a patent.*

Proof. See Appendix. ■

The rational of this proposition is that if the probability of learning decreases sufficiently fast with age and is also decreasing in the return of a patent, then the option value of keeping a patent with low return experiences larger proportional losses than patents with high returns. If patent returns are low, then learning possibilities initially explain a large part of the option value of a patent because learning is more likely for patents with low current returns. Instead, if current returns are high (for instance, upon a potential trade), then learning, although still important, is less weighted in the option value of a patent because newer profitable applications are less likely. Consequently, if learning fades sufficiently fast as age increases, then the option value of a patent with lower current returns experiences larger proportional losses on average than a patent with higher current returns. Thus, given a fixed efficiency gain, patents that were not traded when very young might be traded when they are slightly older. So, $\widehat{g}_a(x,\tau)$ is weakly decreasing in a. However, as time passes the learning effect vanishes and then the process of returns converges to a scenario in which returns depreciate, a particular case of Gibrat's law. Thus, the effect of trading soon surely dominates again. Then, the function $\widehat{g}_a(x,\tau)$ is weakly increasing in a, and the probability of being traded is weakly decreasing in a. Therefore, $\widehat{g}_a(x,\tau)$ is U-shape, which implies that the probability to be traded conditional on survival is hump-shaped with age.

Nevertheless, a learning process by itself does not necessarily imply that $\widehat{g}_a(x,\tau)$ is U-shaped in a. The previous proposition illustrates the necessity of strong assumptions to show that $\widehat{g}_a(x,\tau)$ is not always weakly decreasing in a. Diminishing learning guarantees that the slope of the option value of a patent for the seller is steeper, however it might be the case that it has less steepness than the one of the buyer. In fact, if learning does not diminish fast enough, then the horizon effect dominates and consequently $\widehat{g}_a(x,\tau)$ is increasing in a.

As a matter of fact, we can show that if the process of learning is independent of the age of the patent, then $\widehat{g}_a(x,\tau)$ is weakly increasing in a. This means that the probability to be traded conditional on survival is weakly decreasing. The following proposition shows this result.

Proposition 7 *If $F_{g_a^i}(u^i;z,a)$ is independent of a, then (i) the function $\widehat{g}_a(x,\tau)$ is weakly increasing in a. And (ii) the probability of being traded is weakly decreasing in the age of the patent.*

Proof. See Appendix. ∎

5 Estimation and Identification

In this section, first I introduce a stochastic specification of the model of patent transfers and renewals developed in previous sections. Second, I discuss the estimation strategy and present preliminary estimates of the parameters of the model. Third, I argue the identification of the estimated parameterers.

5.1 Stochastic Specification of the Model

The stochastic specification that is estimated contains 8 parameters. I decided to set the discount factor $\beta = 0.9$ as in Pakes (1986). The rest of the parameters, which are jointly estimated, are those that are contained in the initial distribution of returns, the internal growth of returns and the external growth of returns.

In the model the invention decision is exogenous. So, patents are granted to a fraction of firms. Since patent renewal fees are due by the end of the 4^{th}, 8^{th} and 12^{th} year since the grant date of a patent, I define the age of a patent a as the number of years from its grant date.

Nevertheless, the inventions that a patent, if granted, will eventually protect exist earlier than the grant date of the patent. In fact, the process at which the quality of patent is determined, the internal growth of patent returns, must depend on the number years since the innovation has existed. I have no information specifying the exact that that an innovation was discovered. To deal with this issue, I assume that inventors apply for patents as soon as their innovations are discovered. However, patents granted in a given year were not applied at the same date. For simplicity and following data evidence, I assume that the application period for a patent is at most three years.[20]

In terms of the model, while patent protection is not enforceable until an application is granted, the application period provides information about the quality of an innovation and therefore the future returns of a patent. So, I assume that the realized cash flow of potential patents is zero before a patent is granted, and that the process of internal growth of returns still determines how the quality of a patent evolves during the application process.

The intial quality of an innovation, at the time an inventor applies for a patent, are assumed to be distributed lognormally according to:

$$\log(x_a) \sim F_{IR}(\mu, \sigma_R)$$

where the age is the application year, $a = -2$, and $F_{IR}(.)$ is a normal distribution.

There are two sources that explain how returns evolve over time: internal and external growth of returns. The internal growth of returns occurs within the firm. This process is stochastic and illustrates both the possibility that patent per period returns increase

[20]The application period in the data is calculated to be an average of 2.5 years. In addition, 90% of all issued patents contained in the sample are granted by their third year after the application date.

over time as a result of the discovery of new succesful application, and perhaps returns decrease due to the arrival of competing technologies. The stochastic process is specified as a random variable g_a^i with distribution function $F_{g_a^i}(u^i;a)$. This function is truncated by δ as can be seen as follows.

$$F_{g_a^i}(u^i;a) = \left\{ \begin{array}{lll} 0 & \text{if} & 0 \leq u^i < \delta \\ 1 - \exp(-\frac{u^i}{\sigma_a}) & \text{if} & u^i \geq \delta \end{array} \right\}$$

where the parameter σ_a is defined as $\sigma_a = \phi^{a+2}\sigma$, $a = \{-2, ..., 17\}$. Since it is assumed that $\phi \leq 1$, then σ_a is decreasing as age increases.[21]

Alternatively, the process of internal growth of returns is equivalent to one in which the realized internal growth of returns is the maximum between δ and a draw of a single parameter exponential distribution with support $[0, \infty]$. So the likelihood of arrival of succesful application coincides with the the probability that the internal growth of returns are higher than the depreciation factor, $\Pr(g_a^i > \delta)$, which I define as the *learning likelihood*.

The parameters ϕ and σ determine the learning likelihood as a function of the number of years since the application date of a patent. A small ϕ involves that potential learning possibilities vanish very fast. A low σ implies that the opportunities of learning are not sizeable and that the probability that a patent is valuable is also small.

The growth of returns between firms is modeled as a random variable g^e with an exponential distribution function

$$F_{g^e}(u^e) = 1 - \exp(-\frac{u^e}{\sigma^e})$$

The next period returns of the owner of a patent depend on the internal growth of returns and whether the patent was traded or not.

$$x_{a+1} = g_a^i * z_a$$

Where $z_a \in \{x_a, y_a, 0\}$ is the per period return if the patent is respectively, kept, sold and allowed to expire. And $y_a = g^e x_a$.

The complete stochastic specification contains seven parameters to be estimated, $w = (\mu, \sigma_R, \sigma^e, \tau, \delta, \phi, \sigma)$.

5.2 Estimation Strategy and Estimates

The parameters of the model are structurally estimated using the simulated generalized method of moments. This method involves finding parameters w so that they minimize

[21]Notice that the expectation of the random variables g_a^i is $\mathrm{E}(g_a^i) = \delta + \sigma_a$.

the distance between the empirical moments, defined as those from the data, and the simulated moments generated by the model. The moments generated by the model are simulated because they cannot be solved analytically due to the structure of the model. In particular, my strategy consists in fitting the proportion of patents that are traded conditional on renewal (20 moments) from Figure 1, and the proportion of patents allowed to expire conditional on renewal from the first column in Table 3 (3 moments). To do so I follow the following algorithm.

5.2.1 Estimation Algorithm

The parameters of the model are structurally estimated using the simulated generalized method of moments. A simulating procedure was first applied in a patent renewal model by Pakes (1986), however Pakes' approach used a maximum likelihood estimator. Lanjouw (1998) used the simulated generalized method of moments. In particular, to estimate the parameters of the model I find the simulated minimum distance estimator[22], $\widehat{w}_N$, of the true k parameter vector, w_0[23].

$$\widehat{w}_N = \arg\min_{w} ||h_N - \eta_N(w)||$$

The vector h_N is defined as the sample hazard proportions or empirical moments, and the vector $\eta_N(w)$ as the ones that are simulated. In particular, the vector of empirical moments contains both the hazard to expire, η_N^{ε}, and the hazard to be traded, η_N^{φ}, so $\eta_N(w) = [\eta_N^{\varphi}, \eta_N^{\varepsilon}]$. The vector w_0 is defined as the unique solution to the equation $G(w) = \eta - \eta(w) = 0$, where η is the vector of the true hazard probabilities and $\eta(w)$ are the hazards predicted by the model with parameter vector w.

Using the simulated minimum distance estimator requires simulate the moments $\eta_N(w)$ and minimize the distance between these moments and the empirical counterparts.

The simulated moments $\eta_N(w)$, given a vector of parameters w, are generated in the following way. First, I solve the model recursively and find the cutoff rules $\widehat{x}_a$ and $\widehat{g}_a(x)$ as function of w and c_a. In order to calculate the value function and consequentely the cutoff rules I approximate the double integral that defines the option value of a patent with bilinear approximation methods. Having the cutoff rules, I can construct the hazard to be traded and the hazard to expire for simulated population of patents.

Second, I calculate the simulated moments as the average of the hazards to be traded

[22] The capital letter N denotes the sample size.

[23] Pakes and Pollard (1989) have showed conditions under which $\widehat{w}_N$ converges to w_0, and $\sqrt{n}(\widehat{w}_N - w_0)$ satisfies a central limit theorem.

and the hazards to expire obtained from simulated populations of patents. Each simulated population of patent consists of taking pseudo random draws from the distribution of initial returns F_{IR}, and then I pass each initial patent return through the stochastic process of returns of the model implied by distribution of internal growth of returns $F_{g_d^i}$ and the distribution of external returns F_{g^e}. Finally, I average all simulations and calculate $\eta_N(w)$.[24]

To minimize the distance between the empirical and the simulated moments I use a diagonal weighting matrix with weights depending on the proportional distance between the empirical and simulated moment. In addition, the minimization algorithm is based on simulated annealing methods.[25]

Construction of Simulated Moments Let H_a^s be the distribution function of current patent returns of simulation s for patents of age a, and at the beginning of period a (notice that these patents had been renewed in the past, otherwise their returns would be zero).

$$H_a^s(u) = \Pr\{x_a \leq u, x_{a-1} \geq \widehat{x}_{a-1}, ..., x_{-2} \geq \widehat{x}_{-2}\} \qquad (a = 0, ...17)$$

Patents that are alive up to period $a-1$, must have been renewed in from period 1 to $a-1$ both included, so

$$1 - H_a^s(u) = \Pr\{x_a \geq u, x_{a-1} \geq \widehat{x}_{a-1}, ..., x_{-2} \geq \widehat{x}_{-2}\} \qquad (a = 0, ...17)$$

The model shows that patents with returns higher than $\widehat{x}_a$ will be renewed in period a. Thus, the proportion of patents that survived up to period a, included, is $1 - H_a(\widehat{x}_a)$.

$$1 - H_a^s(\widehat{x}_a) = \Pr\{x_a \geq \widehat{x}_a, x_{a-1} \geq \widehat{x}_{a-1}, ..., x_{-2} \geq \widehat{x}_{-2}\}$$

Now, we can define the proportion of patents that expire in period a conditional on renewal in previous periods. So,the *hazard to expire at age a* is

$$\eta_N^{\varepsilon}(a) = \frac{1}{S}\sum_{s=1}^{S}\frac{H_a^s(\widehat{x}_a) - H_{a-1}^s(\widehat{x}_{a-1})}{H_{a-1}^s(\widehat{x}_{a-1})}$$

The numerator denotes the patents that become expired at age a, and the denominator shows the proportion of all patents still alive after the $(a-1)^{th}$ renewal. And S is the

[24] I run 200 simulations, and in each simulation a do 10,000 draws of patent returns from the initial distribution.

[25] In particular, I use the amebsa and amotsa subroutines as describes in the Numerical Recipes for Fortran at http://www.library.cornell.edu/nr/cbookfpdf.html. See Section 10, Minimization or Maximation of Functions.

number of simulations, that is 200 in my empirical work.

In the case of trades, the proportion of patents that are sold in period a is

$$1 - H_a^s(\widehat{g}_a) = \Pr\{x_a \geq \widehat{g}_a, x_{a-1} \geq \widehat{x}_{a-1}, ..., x_{-2} \geq \widehat{x}_{-2}\}$$

The probability to be traded in period a conditional on renewal, the *hazard to trade at age a* is

$$\eta_N^{\varphi}(a) = \frac{1}{S}\sum_{s=1}^{S}\frac{1 - H_a^s(\widehat{g}_a)}{1 - H_a^s(\widehat{x}_a)}$$

In the U.S. patent system there are three renewal fees, at the 4^{th}, 8^{th}, 12^{th} years since being granted. If the renewal fees are not paid, then patents are immediately expired respectively at the age of 5, 9 and 13 years old.

Finally, $\eta_N^{\varepsilon} = [\eta_N^{\varepsilon}(5), \eta_N^{\varepsilon}(9), \eta_N^{\varepsilon}(13)]$ is the hazard to expire, and $\eta_N^{\varphi} = [\eta_N^{\varphi}(1), ..., \eta_N^{\varphi}(20)]$ is the hazard to be traded.

5.2.2 Estimates

My preliminary estimates are the following.

Table 7: Estimates

Description (Parameter)	**Estimate**
Depreciation factor (δ)	0.79
Learning Factor from Internal Growth of Returns (ϕ)	0.59
Fixed Component from the Mean of the Internal Growth of Returns (σ)	3.87
Mean parameter of the Lognormal Initial Distribution (μ)	7.47
Std. Deviation parameter of the Lognormal Initial Distribution (σ_R)	2.34
Cost of Technology Transfer (τ)	24,454
Mean External Growth of Returns (σ^e)	0.39

5.2.3 Identification

I decided to set the discount factor $\beta = 0.9$ as in Pakes (1986). The rest of the parameters of the model are jointly estimated following the algorithm explained above. The main source of identification of the value of patents is the use of the schedule of renewal fees and their Dollar value together with the functional form assumptions. So, there is no information in the data I currently use in the estimation that identifies an upper bound of the value of patents that are renewed at all renewal dates. The fact that owners allow their patents to expire indicate that the present value of their future expected patent returns is below the one of the renewal fee. Obviously, it is being assumed the fact that owners of patents are willing to pay renewal fees expecting that their future returns will be high enough to compensate these costs.

The parameters of the process of internal growth of returns, which are (δ, ϕ, σ), and the ones from the initial returns, (μ, σ_R), are jointly indentified from the proportion of patents that expire conditional on renewal and the proportion of patents that are traded conditional on renewal. In particular, ϕ is identified from the humpshape form of the transfer rate. A higher ϕ implies that the speed at which learning vanishes is smaller, so the maximum of the transfer rate would shift right and upwards. The parameter δ is identified in part from the expiration rate and the curvature of the transfer rate as patents get closer to their expiration date. Finally, the parameters μ, σ_R are identified from jointly form the expiration rate and also from the level of the transfer rate early in the application period, and σ^e from the level of the transfer rate over the whole life cycle of a patent.

The cost of technology transfer is identified as follows.

Cost of Technology Transfer (τ) The cost of technology transfer is identified by the size of the jumps of the transfer rate that are observed immediately after each renewal date.

The identification strategy is as follows. Let us start considering the case in which the transaction cost is zero. In this case, the model predicts that the transfer rate will be *flat* over the life cycle of a patent except at the renewal rates. Immediately after renewal, the transfer rates increase. In particular, the size of the jumps, which are temporary (i.e. only at the renewal date), depend on the size of the renewal fee c_a. The renewal fee together with the process of external returns determine the curvature of the policy function $\widehat{g}_a(x)$ for any $x < \widehat{x}$ (notice that in this case $\widehat{g}_a(x) = 1$ for all $x > \widehat{x}$). The size of the area delimited by the intersection of $\widehat{g}$ with the upper bound of g^e and $\widehat{x}$ determines the jump of the transfer rate. Finally, as age increases, $\widehat{g}_a(x)$ will shift up, then the area between $\widehat{g}_a(x)$ and the upper bound of its support determines the number of patents traded and consequently the curvature of the transfer rate as a function of age.

6 Estimation Results

This section first discuss the fit and implications of the parameter estimates. Next, it quantifies what are the gains from trade in the market for patents, and the value of a patent.

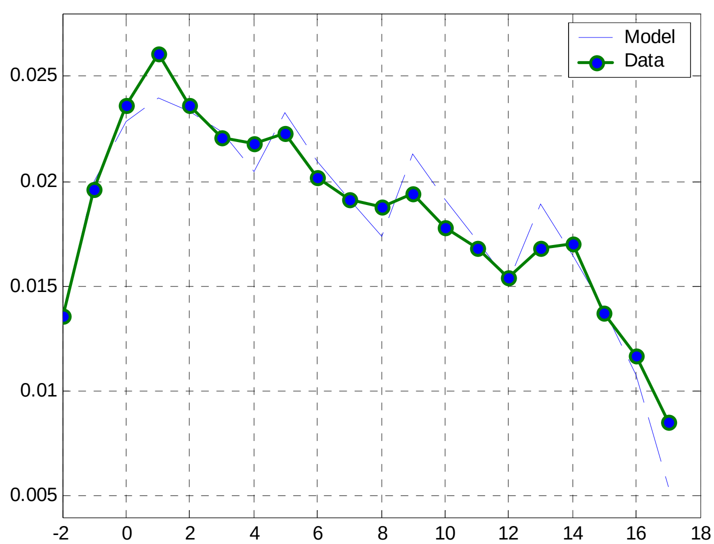
6.1 Fit of the Model

An indicator of how the estimated model fits the data is to compare the empirical moments and the simulated moments from the model. Both the empirical expiration rate and the empirical transfer rate are well tracked by their simulated counterparts as can be seen respectively, in Table 8 and Figure 3. Table 8 shows that both empirical and simulated expiration rate are increasing as a function of the age of the patent. However, the model tends to predict a marginally lower steepness in this moment. Figure 3 shows that the model is able to capture the humpshape form of the transfer rate, including its pick in the grant year of a patent and the sharp decrease of the transfer rate when patent get closer to their expiration date. Nevertheless, the simulation generated by the model tend to overpredict the jumps of the transfer rate immediately after renewal. This feature might be as a result of the timing in which the trading decision occurs in the model. In the model, choices are made at the beginning of the period, while in reality these decisions take place continuosly within a year.

Table 8: Fit of the Expiration Rate

Age	Model	Data
5	0.194	0.181
9	0.276	0.283
12	0.299	0.320

Figure 3: Fit of the Transfer Rate



An additional measure of how the estimated model fits the data is assessing the predictions of the model in other moments that the ones used in the estimation strategy. For instance, we can us the transfer rate conditional on being previously traded or not. In the case the simulated moments generated by the model are remarkably close to the empirical ones, as can be seen in Table 9. The model predicts well the levels and the relationship of the moments with respect to the age of the patent. However, the model tends to overpredict the proportion of traded patents conditional on having being previously traded, and then underpredict the proportion of traded patents conditional on not having been previously traded.

Table 9: Conditional Transfer Rate

		Not Previously Traded		Previously Traded	
		Data	Model	Data	Model
Traded	1	0.024	0.022	0.052	0.058
	4	0.018	0.016	0.042	0.053
	8	0.013	0.010	0.036	0.045
	12	0.010	0.007	0.025	0.034

6.2 Discussion of Estimated Parameters

The parameter estimates are presented in Table 7. The estimates suggest

Initial Returns The parameters μ and σ_R determine the initial distribution of returns. A high σ_R implies a large heterogeneity in quality among patents. A low μ implies that initially the quality or returns of patents are low. For instance, according to the estimated parameters, 20.2% of all patents have an initial per period return below US$ (2003) 250, and the median patent has US$ (2002)1750 in initial per period returns.

Internal Growth of Returns The parameters δ, ϕ, σ determine the process of internal growth of returns. A high σ implies a higher probability of discovering new uses or succesful applications that will increase the returns of the patent. The parameter ϕ measures the speed at which vanishes the learning possibilities. According to the estimates, the learning likelihood of internal growth of patent returns is exhausted by the end of the 5^{th} year of age of a patent, or 8 years since the application date[26]. For instance, the following table uses the estimated parameters of the model to shows the predicted speed at which learning, vanishes. I define *learning* as the Pr $(g_a^i > \delta)$.

[26]Pakes (1986) and Lanjouw (1998) report similar results for the case of German patents.

Table 10: Learning Likelihood and Patent Age

Patent Age	$\Pr(g_a^i > \delta)$
-2	0.82
-1	0.71
0	0.56
1	0.37
2	0.19
3	0.06
4	0.01
5	0.00

The estimate of the depreciation rate shows that when the learning opportunities are exhausted patents returns depreciate fast. For instance, a estimate of $\delta = 0.79$ implies that the per period return of a patent depreciate at the rate of 21% a year. So, a patent with per period returns US$ (2003) 100,000 at age 4 would in average depreciate offering US$ (2003) 11,985 by age 14. This suggests that either competing technologies or imitation erode fast the profits from the protection of intellectual property.

External Growth of Returns The parameter σ^e of an exponential distribution describes the process of potential gross growth of returns. A high σ^e implies both that is more likely that potential buyers are might be more efficient and that the heterogeneity among efficiency gains is also greater. The parameter is estimated to be $\sigma^e = 0.39$. This suggests that the probability that in a given period a potential buyer has larger per period returns is 0.077. However, since the cost of technology transfer is positive, then the decision whether a patent is transferred ultimately depends on both the current per period patent return of its owner and the cost of technology transfer.

Costs of technology transfer The estimate of the cost of technology transfer suggests is US$ (2003) 24,454. This estimate suggests that costs of adopting technologies developed by other firm require important expenditures, perhaps in personel, R&D, or restructuring firm's organization to efficiently use the acquired technology.

Distribution of the Value of Patents In this section the parameter estimates are used to simulate the distributions of the value of patents and show how this distribution evolves as patents become older. These distributions were calculated by generating a cohort of 10,000 patents with their respective initial returns, and consequently following their returns using the estimated process of internal and external growth of returns taking into account whether patent were renewed or traded. The value of a patent is obtained as the discounted present value of its stream of patent returns at a given age.

The simulation show that the distribution of patent returns is very skewed as can be seen by estimated value of patents at their grant date, see Table 11. In particular, the value of the median patent is estimated to be US$ (2003) 27,895, but the bottom 50% of all patents only account for the 3.8% of the total value of all patents in the cohort.

Table 11: Distribution of the Value of Patents at their Grant Date

Percentile	Value (US$ 2003)	Cum % of Total
50	27,895	3.8
75	130,466	23.5
80	207,057	34.8
90	292,465	65.2
98.5	328,350	94.0
99.8	448,098	99.4
Mean value	86,782	

In addition, the distribution of patent returns becomes more skewed as patents become older. The model predicts that low return patents are less likely to be traded and they mainly depreciate at a rate δ. However, higher return patents are more likely to be traded, so their per period returns are less likely to drop. This trading effect make the distribution of patent value to be more skewed than in the grant date.

Table 12: Distribution of the Value of Patents at Age 4

Percentile	Value (US$ 2003)	Cum % of Total
50	10,823	2.2
75	104,841	15.9
80	164,301	24.3
90	266,148	58.3
98.5	318,168	93.5
99.8	433,376	98.8
Mean value	66,769	

The next section focuses on the gains from trade in the market for patents.

6.3 Gains from Trade

Patents are traded because some firms are more productive than others in the use of a given patent. So, upon the transfer of a patent, gains from trade are realized. In the model, I define the gains from trade as the value of patent rights generated by the ability to transfer patents between firms.

To calculate the value of the gains from trade I run the following expirement. First, I use the above estimates and run simulations to obtain the average value of a patent at its grant date. I find that the average present value of a patent is US$ (2003) 86,782. Next I shutdown the market for patents by setting the cost of technology transfer high enough so that no patent is traded. Consequently, I find that the average value of a patent by its grant date is US$ (2003) 82,585. The gains from trade is the ratio between the two values, that is 5.08%[27].

This estimate is a lower bound of the gains from trade in the market for intellectual property. Patents might be licensed too and that also results in gains from trade, however there is not systematic data on licensing because these transactions tend to be private agreements between firms.

A statistic that is also of interest is to quantify what are the *gross* gains from trade, defined as those when the cost of technology transfer equal to zero. In this case, I find that the gross gains from trade are 9.05%. This statistic is an upper bound of how much can be achieved by using policy addressed at reducing the cost of technology transfer to promote the transfer of technology. In particular, I show that reducing the cost of technology transfer by 50% implies that the gains from trade of the market would increase from 5.08% to 6.58%, this is 1.5 percentage points more or a 29% increased.

7 Conclusion

This paper is the first systematic work that uses data on the transfer of patents to shed light on the workings of the market for intellectual property. It is a significant advancement because the previous literature on patents and intellectual property transfer have been hampered by a lack of systematic data on how intellectual property assets are traded. This study attains two objectives. First, it presents the stylized facts about the transfer of patents. Second, it develops and analyzes a model of patent trades that explains the stylized facts. The paper finds evidence that the market for trading patents increases efficiency, with patents rising in value as they are better matched with firms.

[27]The gains from trade are net of taxes and the cost of technology transfer.

The data presented in this paper has been collected using recorded transfer of patents at the USPTO. The dataset compiled is a panel of patents with their histories of renewal and trading decisions over their life cycle. The paper moves on to present the stylized facts observed in the data. First, patents are traded often and traded patents are of higher quality. In particular, 20% of all patents granted to small innovators are traded at least once over their life cycle. These patents represent 35% of all ex post citations received. Second, better patents are more likely to be traded. Logistical analysis shows that an extra citation received by a patent increases the log of the probability of being traded by 0.013 units. Third, the transfer rate varies over the life cycle of patents. It is hump-shaped as a function of the age of the patent: (i) it decreases from its grant date with the exception of the renewal dates. Immediately after renewal, the transfer rate increases. (ii) it increases during the application period. Finally, patents that have been previously traded, especially the ones recently traded, are more likely to be retraded and less likely to expire.

The stylized facts and intrinsic characteristics of patents motivate a dynamic model of patent trades with costly technology transfer. In the model, there are two mechanisms. First, the cost of technology transfer creates a selection effect so that better patents are more likely to be traded. This explains the discontinuous increase of the transfer rate after the renewal decision, and the evidence that traded patents are more likely to be traded and less likely to expire. Second, there exists a horizon effect that explains that the transfer rate decreases as the patent gets closer to its expiration date. This is because the shorter horizon implies less time to amortize the cost of technology transfer. This accounts for the observed decreasing transfer rate over the life cycle of patents.

The parameters of the model are estimated using the simulated general method of moments. I find preliminary estimates that show that the market for patents add 5.08% to the value of a patent. This number is a lower bound of the gains from trade in the market for intellectual property because licensing opportunities have not been considered in this accounting. In the near future, I plan to use patent citations received as a proxy of patent value in order to reestimate the parameters of the model.

This work opens new avenues of research. Perhaps most interesting would be to study the sources of innovation and to characterize who are the buyers and sellers of technology. In particular, to trace the flow of technology transfer, and to analyze whether small firms specialize in innovating and then selling their inventions to larger firms, which might have a comparative advantage in their management. Second, to evaluate to what extent the move toward higher protection of patent rights that occurred in the 1980's has facilitated specialization, and consequently trade in patents. Lastly, this work can also be extended to examine alternatives to promote innovation such as lower taxation on intellectual property transfer.

These questions have not been previously addressed empirically due to a lack of data on how intellectual property assets are traded.

References

[1] Acs, Z. and Audretsch, D. (1988), "Innovation in Large and Small Firms: An Empirical Analysis." *The American Economic Review*, Vol. 78, No. 4, 678-690.

[2] Arora, Fosfuri and Gambardella (2002), "Market for Technology: The Economics of Innovation and Corporate Strategy." The MIT Press. Cambridge, Massacheusetts.

[3] Arrow, K. (1962). "Economic Welfare and the Allocation of Resources for Invention." In the Rate and Direction of Inventive Activity. Special conference series NBER. Princeton, NJ: Princeton Univ. Press.

[4] Arrow, K. (1962A). "Comment on Case Studies. In The Rate and Direction of Inventive Activity. Special conference series NBER. Princeton, NJ: Princeton Univ. Press.

[5] Arrow, K. (1983). "Innovation in Large and Small Firms." In Entrepreneurship, edited by Joshua Ronen. Lexington, Mass.: Lexington Books.

[6] BLS (1979). *The Impact of Research and Development on Productivity Growth*. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics, Bulletin 2331.

[7] Griliches, Z. (1979). "Issues in Assesing the Contribution of R&D to Productivity Growth." *Bell Journal of Economics*, 10, 92-116.

[8] Griliches, Z. (1992). "The Search of R&D Spillovers." *Scandinavian Journal of Economics*, 94, 29-47.

[9] Hall, B., Jaffe, A. and Tratjenberg M. (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper 8498.

[10] Hall, B., Jaffe, A. and Tratjenberg M. (2004). "Market Value and Patents Citations." *Rand Journal of Economics*, forthcoming 2004.

[11] Hall, B., Griliches, Z., Hausman, J. (1986). "Patents and R and D: Is There a Lag?" *International Economic Review*, 27, 265-283.

[12] Jaffe, A., Henderson R. and Tratjenberg M. (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*. Vol. CVIII, issue 3, p. 577.

[13] Holmes, T. and Schmitz, J. (1990). "A Theory of Entrepreneurship and Its Application to the Study of Business Transfers." *Journal of Political Economy*, vol. 98, no. 2.

[14] Holmes, T. and Schmitz, J. (1995). "On the Turnover of Business Firms and Business Managers." *Journal of Political Economy*, vol 103, no. 5.

[15] Lanjouw, J. O. (1998). "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value". *Review of Economic Studies*, vol 65.

[16] Lanjouw, J. O., Pakes, A. and Putnam, J. (1998), "How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data." *Journal of Industrial Economics*.

[17] Lamoreaux, N. and Sokoloff, K. (1997). "Inventors, Firms, and the Market for Technology: U.S. Manufacturing in the Late Nineteenth and Early Twentieth Centuries." NBER Working Paper H0098.

[18] Lamoreaux, N. and Sokoloff, K. (1999). "Inventive Activity and the Market for Technology in the United States, 1840-1920." NBER Working Paper 7107.

[19] Pakes, A. (1986). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica*, vol. 54, no. 4

[20] Pakes, A. and Griliches, Z. (1984). "Patents and R&D at the Firm Level: A First Look." In *R&D, Patents and Productivity*, ed. by Z. Griliches. Chicago: University of Chicago Press.

[21] Serrano, C. (2004). "Measuring the Transfer of Patents." Manuscript.

[22] Teece, D. J. (1977). "Technology Transfer by Multinational Firms: The Resource Cost of Transferring Technological Know-how." *Economic Journal*, 87: 242-261.

[23] Tratjenberg M., Henderson R. and Jaffe A. (1997). "University versus Corporate Patents: A Window on the Basicness of Invention." *Economics of Innovation and New Technology*, 5, 19-50.

[24] Williamson O.E. (1991). "Comparative Economic Organization -The Analysis of Discrete Structural Alternatives." Administrative Science Quarterly 36 (4): 269-296.

[25] WIPO (2003). "Intellectual Property: A Power Tool for Economic Growth." World Intellectual Property Organization (WIPO)

Appendix 1: Summary Statistics.

Table 5: Patents Traded Based on Non-Truncated Cohorts (i.e., 1983, 1984 and 1985)

	Small innovators		All innovators	
Age	Counts	Transfer Rate	Counts	Transfer Rate
0	1049	2.47	1965	1.45
1	860	2.27	2116	1.81
2	981	2.35	2646	2.13
3	947	2.21	3145	2.15
4	927	2.15	3029	1.95
5	779	2.15	2431	1.63
6	670	1.81	1852	1.33
7	628	1.76	1657	1.27
8	572	1.70	1473	1.20
9	449	1.90	1244	1.37
10	460	1.91	1294	1.48
11	398	1.72	1303	1.43
12	397	1.65	1142	1.29
13	293	2.02	868	1.61
14	339	2.15	898	1.66
15	263	1.63	850	1.40
16	223	1.05	791	0.98
17	128	0.60	454	0.61

Table 6: Patents Granted and Mean Citations Received For Incomplete and Complete (spells) Cohorts

	All innovators		Small innovators	
	Counts	Mean Citations	Counts	Mean Citations
1983	33190	7.15	9980	7.02
1984	50173	7.30	15274	6.99
1985	56178	7.39	17012	7.00
1986	55599	7.44	17798	7.13
1987	65836	7.63	20397	7.35
1988	62313	7.39	20037	7.03
1989	76095	6.96	23246	6.54
1990	71603	6.63	22354	6.28
1991	76702	6.16	23106	5.88
1992	78520	5.73	23139	5.32
1993	80202	5.13	23551	4.74
1994	82801	4.45	24489	4.00
1995	82705	3.50	24638	3.01
1996	90200	2.48	25630	2.07
1997	93121	1.35	26325	1.11
1998	123714	0.50	32037	0.40
1999	129553	0.04	33354	0.03
2000	134047	na	36390	na
2001	143254	na	36656	na
2002	145811	na	36198	na

Appendix 2: Proofs

Proof of Lemma 1:

The proof is an extension based on Pakes (1986) results.

Proof of Proposition 2:

For convenience of notation, let us rewrite the value function $V(a,x,y)$ as $V_a(x,y)$. I want to show that $\widehat{g}_a(x,\tau)$ is weakly decreasing with respect x. I have divided the proof into two parts. The first one studies the case where $\widehat{g}_a(x,\tau)$ is defined as the external growth of returns that makes a firm indifferent between selling and allow the patent to expire. This result is straightforward. Formally the function $\widehat{g}_a(x,\tau)$ is defined as

$$V(a, \widehat{g}_a(x,\tau)x) = 0$$

From Lemma 1 we know that the value function is weakly increasing in per period returns of a patent. Let us suppose that x increases, it must be the case that $\widehat{g}_a(x,\tau)$ is decreasing with x to keep the equality holding.

The second part of the proof analyzes the subtle case in which $\widehat{g}_a(x,\tau)$ is defined as the external growth of returns that makes a firm indifferent between selling and keep the patent. Consider assumption G, that states the internal growth of returns is independent of the level. Patent returns evolve over time according to

$$\begin{aligned} g_a^i &= \frac{x'}{z} \\ g^e &= \frac{y'}{x'} = \frac{y'}{g^i z} \end{aligned}$$

where in the case of assumption G, the joint density function of the random variables g_a^i and g^e, defined as $f_a(g^x, g^y)$ depends upon a but not z.

The argument of the proof is by induction on the age of the patent. Let us start by considering the last period, $a = L$.

$$\begin{aligned} \left[\hat{g}_L(x,\tau) - 1\right] x &= \tau \\ \hat{g}_L(x,\tau) &= \frac{\tau}{x} + 1 \end{aligned}$$

which is decreasing in x.

Now, assume true for $a' > a$, I will show is also true for a.

The decision of whether to sell or keep relies on the following expression.

$$\begin{aligned}
& \widetilde{V}_a^S(x,y) - \widetilde{V}_a^K(x,y) \\
= \;& V_a(y) - \tau - V_a(x) \\
= \;& V_a(g_a^y x) - \tau - V_a(x) \\
= \;& (g_a^y - 1)\, x - \tau \\
& + \beta \left[E\left[\widetilde{V}_{a+1}(x', y') | a, g_a^y x \right] - E\left[\widetilde{V}_{a+1}(x', y') | a, x \right] \right]
\end{aligned}$$

It is sufficient to show that the above is weakly increasing in x. The first term is increasing in x. Now look at the second

$$\begin{aligned}
& E\left[\widetilde{V}_{a+1}(x', y') | a, g_a^e x \right] - E\left[\widetilde{V}_{a+1}(x', y') | a, x \right] \\
= \;& \int_{g_{a+1}^x} \int_{g_{a+1}^y} \left[\begin{array}{c} \widetilde{V}_{a+1}(g_{a+1}^i g_a^e x, g_{a+1}^e g_{a+1}^i g_a^e x) - \\ \widetilde{V}_{a+1}(g_{a+1}^i x, g_{a+1}^e g_{a+1}^i x) \end{array} \right] f_a(g^i, g^e) dg_{a+1}^i dg_{a+1}^e
\end{aligned}$$

For general a, recall that as an induction hypothesis we assumed

$$\widetilde{V}_{a+1}^S(x,y) - \widetilde{V}_{a+1}^K(x,y)$$

was weakly increasing in x. It suffices (given conditions on the joint density funtion of the growth of returns) to prove that the interior of the double integral is increasing in x.

$$\widetilde{V}_{a+1}(g_{a+1}^i g_a^e x, g_{a+1}^e g_{a+1}^i g_a^e x) - \widetilde{V}_{a+1}(g_{a+1}^i x, g_{a+1}^e g_{a+1}^i x)$$

There are four cases to study.

1. in K region with $(g_{a+1}^x x, g_{a+1}^e g_{a+1}^i x)$, in K region with $(g_{a+1}^i g_a^e x, g_{a+1}^e g_{a+1}^i g_a^e x)$

$$\begin{aligned}
& \widetilde{V}_{a+1}(g_{a+1}^i g_a^e x, g_{a+1}^e g_{a+1}^i g_a^e x) - \widetilde{V}_{a+1}(g_{a+1}^i x, g_{a+1}^e g_{a+1}^i x) \\
= \;& V_{a+1}(g_{a+1}^i g_a^e x) - V_{a+1}(g_{a+1}^i x)
\end{aligned}$$

Define $\lambda = g_a^e$. It suffices to show that the above is weakly increasing in x. The above is an increasing transformation of the induction hypotheses because λ is independent of $g_{a+1}^i x$. Then the above expression is increasing in x.

2. in K region with $(g_{a+1}^i x, g_{a+1}^e g_{a+1}^i x)$, in S region with $(g_{a+1}^i g_a^e x, g_{a+1}^e g_{a+1}^i g_a^e x)$

$$\begin{aligned}
& \widetilde{V}_{a+1}(g_{a+1}^i g_a^e x, g_{a+1}^e g_{a+1}^i g_a^e x) - \widetilde{V}_{a+1}(g_{a+1}^i x, g_{a+1}^e g_{a+1}^i x) \\
= \;& V_{a+1}(g_{a+1}^e g_{a+1}^i g_a^e x) - \tau - V_{a+1}\left(g_{a+1}^i x\right)
\end{aligned}$$

I want to show that the above in increasing in x. Let $\lambda = g^e_{a+1}g^e_a$. Since λ is independent of $g^i_{a+1}x$ and by the previous induction argument, then the above is weakly increasing in x.

Similarly, it can be shown for the remaining two cases.

3. in E region with $(g^i_{a+1}x, g^e_{a+1}g^i_{a+1}x)$, in E region with $(g^i_{a+1}g^e_a x, g^e_{a+1}g^i_{a+1}g^e_a x)$
4. in S region with $(g^i_{a+1}x, g^e_{a+1}g^i_{a+1}x)$, in S region with $(g^i_{a+1}g^e_a x, g^e_{a+1}g^i_{a+1}g^e_a x)$

This completes the proof.

Proof of Proposition 3:

I want to show that $\widehat{g}_a(x,\tau)$ is weakly increasing in a. Consider assumption G, that states the internal growth of returns is independent of the level. Patent returns evolve over time according to

$$
\begin{aligned}
g^i_a &= \frac{x'}{z} \\
g^e_a &= \frac{y'}{x'} = \frac{y'}{g^i_a z}
\end{aligned}
$$

where in the case of assumption G, the joint density function of the random variables g^i_a and g^e, defined as $f_a(g^x, g^y)$ depends upon a but not z.

For convenience of notation, let us rewrite the value function $V(a,x,y)$ as $V_a(x,y)$. To prove the proposition, it suffices to show that the difference $\widetilde{V}^S_a(x,y) - \widetilde{V}^K_a(x,y)$ is decreasing in a, so that $\hat{g}_a(x,\tau)$ is increasing in a.

$$
\begin{aligned}
& \widetilde{V}^S_a(x,y) - \widetilde{V}^K_a(x,y) \\
= \; & V_a(y) - \tau - V_a(x) \\
= \; & V_a(g^e_a x) - \tau - V_a(x) \\
= \; & (g^e_a - 1)\,x - \tau \\
& + \beta\left[E\left[\widetilde{V}_{a+1}(x',y')|a, g^e_a x\right] - E\left[\widetilde{V}_{a+1}(x',y')|a, x\right]\right]
\end{aligned}
$$

The argument of the proof is by induction. First, we start in the case of $a = L$ and $L-1$, that is the last and penultimate period of life of a patent

$$
\begin{aligned}
& \widetilde{V}^S_L(x,y) - \widetilde{V}^K_L(x,y) \\
= \; & x g^e_L - \tau - x \\
= \; & x(g^e_L - 1) - \tau
\end{aligned}
$$

For $a = L - 1$, it is

$$\begin{aligned} & \widetilde{V}^S_{L-1}(x, y) - \widetilde{V}^K_{L-1}(x, y) \\ = \;& x(g^e_{L-1} - 1) - \tau + \beta[E\left[\widetilde{V}_L(x', y')|L-1, g^e_{L-1}x\right] - E\left[\widetilde{V}_L(x', y')|L-1, x\right]] \end{aligned}$$

There are three cases to study. The first case is the one in which the patent was kept in period $L-1$ as well as in period L. The second case considers the possibility of sale in period $L-1$ and being kept in period L. Finally, the third one analyzes the case that the patent is sold in both periods.

(1) K - K

$$\begin{aligned} & x(g^e_{L-1} - 1) - \tau + \beta[E[\widetilde{V}^K_L(x', y')|L-1, g^y_{L-1}x] - E\left[\widetilde{V}^K_L(x', y')|L-1, x\right] \\ = \;& x(g^e_{L-1} - 1) - \tau + \beta[E[V_L(g^i_L g^e_{L-1}x)] - E[V_L(g^i_L x)]] \\ = \;& x(g^e_{L-1} - 1) - \tau + \beta \int [v_L(g^i_L g^e_{L-1}x) - v_L(g^i_L x)] f_L(g^i, g^e) dg^i \end{aligned}$$

Taking as given x, g^e_{L-1}, and g^e_L, $\widetilde{V}^S_a(x, y) - \widetilde{V}^K_a(x, y)$ is decreasing in a because the integral is larger or equal than zero.

Similarly we can also show it for the next two cases.

For a general a,we know that $\left[E\left[\widetilde{V}_a(x', y')|a-1, g^e_{a-1}x\right] - E\left[\widetilde{V}_a(x', y')|a-1, x\right]\right] = 0$ for $a = L$. So, we have to show that

$$\{E\left[\widetilde{V}_{a+1}(x', y')|a, g^e_a x\right] - E\left[\widetilde{V}_{a+1}(x', y')|a, x\right]\} \to 0 \text{ uniformly}$$

Since we know that

$$\begin{aligned} E\left[\widetilde{V}_a(x', y')|a-1, g^e_{a-1}x\right] &\to 0 \text{ as } a \text{ approaches } L \\ E\left[\widetilde{V}_a(x', y')|a-1, x\right] &\to 0 \text{ as } a \text{ approaches } L \end{aligned}$$

and by proposition 1 we know that given a and $g^e > 1$ if a trade takes place

$$E\left[\widetilde{V}_a(x', y')|a-1, g^e_{a-1}x\right] \geq E\left[\widetilde{V}_a(x', y')|a-1, x\right]$$

Therefore $\widehat{g}_a(x, \tau)$ increases in a because $f_a(g^i, g^e)$ depends upon a but not on the return of the patent, z.

Finally, I show that the probability of being traded conditional on survival is weakly decreasing as a function of age. Since it is assumed that the process of arrival of offers, $F_{g^e}(u^e)$, is independent of age, and we have already shown that $g_a(x,\tau)$ is weakly increasing in a, then the probability of being traded is weakly decreasing in a.

Proof of Lemma 4:

We want to show that $\left|\frac{\Delta V(a+1,x',y'|z)}{\Delta a}\right| \leq Q$, which rearranging is $|V(a+1,x',y'|z) - V(a,x',y'|z)| \leq Q$. It is sufficient to show that $(V(a,x',y'|z) - V(a+1,x',y'|z)) \leq Q$.

Let M be the maximum value that a return can achieve. It is finite because the distribution of the growth of internal and external returns are bounded above. Then

$$\begin{aligned} V(a,x',y'|z) &= z + \beta E(z') + \beta^2 E(z'') + ... + \beta^{17-a} E(.) \\ &\leq M + \beta M + \beta^2 M + ... + \beta^{17-a} M \\ &= M \frac{(1-\beta^{18-a})}{1-\beta} \end{aligned}$$

Let m be the minimum value that a return can achieve (i.e., considering that the process of internal growth of returns is bounded below by δ). So, we can show that

$$\begin{aligned} V(a+1,x',y'|z) &= z + \beta E(z') + \beta^2 E(z'') + ... + \beta^{17-a-1} E(.) \\ &\leq m + \beta m + \beta^2 m + ... + \beta^{17-a-1} m \\ &= m \frac{(1-\beta^{18-a-1})}{1-\beta} \end{aligned}$$

Therefore, it suffices for Q to be such that $M\frac{(1-\beta^{18-a})}{1-\beta} - m\frac{(1-\beta^{18-a-1})}{1-\beta} \leq Q$.

Proof of Proposition 5:

Given any patent return z_1 and z_2 such that $z_2 > z_1$, I want to show that

$$\frac{\Delta[V(a,x',y'|z_2) - V(a,x',y'|z_1)]}{\Delta a} = \left\{ \begin{array}{ll} > 0 & \text{if} \quad a < a^* \\ \leq 0 & \text{if} \quad a \geq a^* \end{array} \right\}$$

It is sufficient to examine the sign of the relationship between the slopes of the option value of a patent, that is the following expression.

$$\left(\frac{\Delta EV(a+1,x',y'|z_2)}{\Delta a} - \frac{\Delta EV(a+1,x',y'|z_1)}{\Delta a} \right)$$

Without loss of generality, I focus on the case in which the holder of the patent chooses to keep it. Similarly, we can also show the result for the case in which the patent is sold.

$$
\begin{aligned}
& \left(\frac{\Delta EV(a+1,x'|z_2)}{\Delta a} - \frac{\Delta EV(a+1,x',z_1)}{\Delta a}\right) \\
= \quad & \frac{\Delta V(a+1,\delta z_2)}{\Delta a} F_{g_i}(\delta;a,z_2) - \frac{\Delta V(a+1,\delta z_1)}{\Delta a} F_{g_i}(\delta;a,z_1) \\
& + V(a+1,\delta z_2)\frac{\Delta F_{g_i}(\delta;a,z_2)}{\Delta a} - V(a+1,\delta z_1)\frac{\Delta F_{g_i}(\delta;a,z_1)}{\Delta a} \\
& + \int_{\delta^+}^{\infty}[\frac{\Delta V(a+1,u^i z_2|z_2)}{\Delta a} f_{g_i}(u^i;a,z_2) - \frac{\Delta V(a+1,u^i z_1|z_1)}{\Delta a} f_{g_i}(u^i;a,z_1)]\mathrm{d}u^i \\
& + \int_{\delta^+}^{\infty}[V(a+1,u^i z_2|z_2)\frac{\Delta(f_{g_i}(u^i;a,z_2)}{\Delta a} - V(a+1,u^i z_1|z_1)\frac{\Delta(f_{g_i}(u^i;a,z_1)}{\Delta a}]\mathrm{d}u^i
\end{aligned}
$$

where f_{g_i} is the density function of F_{g_i}.

The first term is always negative. To show this, first notice that $\frac{\Delta V(a+1,\delta z_i)}{\Delta a} < 0$ for $i \in \{1,2\}$. It was assumed that $F_{g_i}(\delta;a,z_2) > F_{g_i}(\delta;a,z_1)$. Using the particular case in which F_{g_i} is independent of z we can show that $\frac{\Delta V(a+1,\delta z_2)}{\Delta a} < \frac{\Delta V(a+1,\delta z_1)}{\Delta a}$ always holds. Therefore, the first term is negative.

We can also show that the sign of the second term is positive. By Lemma 1 we know that the value function is increasing in the patent per period returns, so $V(a+1,\delta z_2) > V(a+1,\delta z_1)$. Also we have assumed that the probability of "no learning" (i.e., $g_a^i = \delta$) is increasing in a, that is equivalent to $\frac{\Delta F_{g_i}(\delta;a,z_2)}{\Delta a} > \frac{\Delta F_{g_i}(\delta;a,z_1)}{\Delta a}$. The term is clearly positive.

The sign of the third term is ambiguos. On the one hand, we know that $\frac{\Delta V(a+1,u^i z_2)}{\Delta a} < \frac{\Delta V(a+1,u^i z_1)}{\Delta a}$, however $f_{g_i}(u^i;a,z_2) < f_{g_i}(u^i;a,z_1)$. This result is not surprising. For instance, in the case in which the the process F_{g_i} is independent of z the equivalent of this term would be always negative. Introducing the assumption for which learning is less likely for patents with higher returns makes the expected value of the potential buyer smaller than in the case of independence.

Since we assumed that $\frac{\Delta(f_{g_i}(u^i;a,z_2))}{\Delta a} > \frac{\Delta(f_{g_i}(u^i;a,z_1))}{\Delta a}$, then the sign of the fourth term is positive.

Next, we have to show that for sufficiently small a, the difference between the slopes of the option value for the potential buyer and seller, $\frac{\Delta EV(a+1,x'|z_2)}{\Delta a} - \frac{\Delta EV(a+1,x',z_1)}{\Delta a}$, increases with age. To do so, it is sufficient to show that the second and fourth terms are larger than the first and third term. The strategy is to construct a bound for the term $\frac{\Delta V(a+1,x',y'|z)}{\Delta a}$. By Lemma 2 we know that $\left|\frac{\Delta V(a+1,x',y'|z)}{\Delta a}\right| \leq Q$, where Q is a number sufficiently small and positive. Therefore, this bound permit us to make the first and third term small enough to compare to the other two terms. There exists a a^* such that $a < a^* \leq \overline{a}$ that the second and fourth term dominate the first and the third one. If $a \geq a^*$, then the last three terms become monotonically small converging to zero when learning vanishes at age $a = \overline{a}$ (i.e., the probability of learning is nul) (i.e., this is a particular case of Gibrat's law, see Proposition 2).

Proof of Proposition 6:

The argument of the proof is as follows. It is assumed that the returns due to internal growth of returns are subject to first order stochastically dominance in z. In other words, the higher today's return is, the more likely it is that the return of tomorrow will be high. Now, consider the problem of whether to sell a patent in period L, that is the last period of life of a patent. A patent is sold if $y \geq x + \tau$. Consider now a period before the last, which is $L - 1$. A patent is now sold if the following condition holds

$$y + \beta E(y'|y) \geq x + \beta E(x'|x) + \tau$$

Rearranging this condition, we obtain that

$$y \geq x + \beta[E(x'|x) - E(y'|y)] + \tau$$

However $\beta[E(x'|x) - E(y'|y)] \leq 0$ because of first order stochastically dominance. Then, it must be the case that $\widehat{g}_{L-1} \leq \widehat{g}_L$. The proof can be extended backwards by any number of finite periods. Therefore, $\widehat{g}_a(.)$ is weakly increasing in a.

Proof of Proposition 7:

I want to show that $\widehat{x}_a(\tau)$ is increasing in a. First I show that for any age a there exists a unique $\widehat{x}_a(\tau)$ for which firms are indifferent between keeping and discontinuing a patent. So, $\widehat{x}_a(\tau)$ is defined as $V(a, \widehat{x}_a(\tau)) = 0$. It is assumed that the schedule of renewal fees c_a, when positive, is increasing with age. It is also assumed in the paper that the probability that tomorrow's return is larger than a given number u is weakly decreasing with age. Given that, Lemma 1 demonstrates that the option value function of a patent, $E\widetilde{V}(a + 1, x', y'|a, z)$ is weakly decreasing in a. Therefore, the returns that make a firm indifferent between keeping and discontinuing, $\widehat{x}_a(\tau)$, must be increasing as age increases.