ESTIMATING THE GAINS FROM TRADE IN THE MARKET FOR INNOVATION:
EVIDENCE FROM THE TRANSFER OF PATENTS

Carlos J. Serrano

Estimating the Gains from Trade in the Market for Innovation: Evidence from the Transfer
of Patents
Carlos J. Serrano
NBER Working Paper No. 17304
August 2011
JEL No. L24,O32,O34

ABSTRACT

The "market for innovation" — the sale and licensing of patents — is an often discussed source of
incentives to invest in R&D. This article presents and estimates a model of the transfer and renewal
of patents that, under some assumptions, allows us to quantify the gains resulting from the transfer
of patents in the market for innovation. The gains from trade measure the benefits of reallocating the
ownership of a patent from the original inventor to a new owner for whom the patent has a higher
value. In addition, we study the effect that lowering the costs of technology transfer has on the proportion
of patents traded and the gains from trade.

Carlos J. Serrano
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
CANADA
and NBER
carlos.serrano@utoronto.ca

1 Introduction

The "market for innovation" – the sale and licensing of patents– is an often discussed source of incentives to invest in R&D. The gains that original inventors can obtain from this market — and the corresponding incentives to invest in R&D — are greater the greater are the extra-profits that non-inventors can generate from the innovation. Previous research scholars have associated these gains from trade to comparative advantages in manufacturing and advertising (Teece, 1986; Arora and Ceccagnoli, 2006), to vertical specialization (Teece, 1986; Arora and Gambardella, 1994; Arora, Fosfuri and Gambardella, 2001), and to the impossibility to control ex-ante the exact nature and scope of utilization of innovations (Rosenberg, 1996). Moreover, transactions in patent rights are also important because they can increase the nature of competition between the small firms, that lack the capacity of large scale development, and the large firms, that may have a comparative advantage in the development and commercialization process (Gans, Hsu and Stern, 2002).

Despite the potential benefits associated with markets for innovation, little empirical systematic evidence exists on the importance of these benefits for firms. In this paper, we provide evidence on the gains from trade in the market for innovation. The magnitude of these benefits has been the subject of many recent debates about patent policy reform and the importance that small firms play in the economy.[1] If these benefits were significant, they could potentially offset traditional patent policy concerns about increased market power and inefficiencies associated with the granting of monopoly rights. For small firms, being innovations often their critical assets, the ability to license and sell their patents may be an important source of their incentives to invest in R&D, especially because of the lack of capacity of large scale development and commercialization.

Empirical studies quantifying the benefits from the market for innovation have traditionally been hampered by a lack of systematic data on the transactions in patent rights (Arora, 1997; Arora, Fosfuri and Gambardella, 2001; Robbins, 2009). Our analysis focuses on the sale of patents. From an empirical point of view, an interesting feature of the sale of patents —that distinguishes it from the licensing of patents— is the fact that the sale of patents is most often publicly recorded because of the legal requirement that all patent sales have to be filed with the United States Patent and Trademark Office (USPTO) in order to be legally binding (Dykeman and Kopko, 2004). This property of patent sales allows us to link these transactions to patent numbers, which can then be merged to the basic patent data that others have used: the payment of renewal fees, patent citations, the patent's issue date, a patentee identifier, etc. (Griliches, 1992; Pakes and

---

[1]See report on the "The Evolving IP Marketplace: Aligning Patent Notice and Remedies with Competition" by the Federal Trade Commission (March, 2011).

Schankerman, 1984; Hall, Jaffe and Tratjenberg, 2001). Moreover, the market for the sale and acquisition of patents is increasingly popular for firms. For instance, large firms such as IBM and Dell have recently made available for sale hundreds of their patents. At the same time, several patent auctions, patent exchanges, and patent brokers have emerged in recent years.[2] But, the sale of patents is not important just solely for large firms. For small firms the rate of the sale of patents is higher than for the large firms, and this number is much higher when weighted by patent quality (Serrano, 2010).

The information on the sale and renewal of patents allows us, under some assumptions, to recover the gains from trade and the distribution of the value of holding patents. To do so, we develop and estimate a model of patent sales and renewals. The starting point of our model is Schankerman and Pakes (1986) and Pakes (1986). They examine the problem of a patent owner deciding in each time period whether or not to pay a renewal fee, and thereby to extend the life of a patent, in a context with heterogeneity in the economic value of inventions. The distinct element of our theory is the introduction of the possibility of the arrival of opportunities for surplus-enhancing transfers, which may lead to a new owner for whom the patent has a higher value. However, to transfer the patent to a new owner involves a cost of technology transfer. The gains from trade measure the net benefits of reallocating the ownership of a patent from the original inventor to a new owner. Therefore, whereas Pakes and Schankerman's framework has one margin —whether the patent owner should pay the fee for renewing the patent— our model has a second margin —whether the cost of technology transfer should be covered to reallocate the patent to an alternative owner.

In our empirical analysis, we will focus on patents applied for and granted to small firms. We operationalize this focus on small firms by restricting attention to firms with no more than 500 employees as of the application date of the patent. There is a number of reasons that support our focus on small firms. The first one is that in the policy arena there exists special interest in the importance that the market for innovation plays on the incentives to invest in R&D by small firms. Another reason is that by focusing on small firms the economic forces that we highlight will be more salient than in transactions involving the patents from larger firms.[3] At the same time, the focus on small firms allows us to parallel our empirical analysis with the existing theory on patent renewals, in which the decision making is at the patent level. Furthermore, small firms are interesting in their own right, given the importance they play in the innovation process (Arrow,

---

[2]Examples of patent brokers and market intermediaries are Inflexion Point, IPotential, Ocean Tomo, ICAP.

[3]Serrano (2010) reports that patents granted to large corporations may not only be traded for the technology that they represent, but also as a result of large acquisitions that may be pursued to increase the buyer's market share in a particular product market, etc.

1983 and Acs and Audretsch, 1998).

The parameters of the model are estimated using the simulated generalized method of moments (McFadden, 1989 and Pakes and Pollard, 1989). The empirical moments we use to identify these parameters describe both the patent life cycle properties of the trading and expiring decisions of patent owners. The first set of moments that refer to the trading decision: (1) the probability that an active patent is traded at alternative ages conditional on having been previously traded; and (2) the probability that an active patent is traded at alternative ages conditional on having been previously untraded. The second set of moments relates to the expiring decision and how does this decision interplays with past trading history of patents: (3) The probability that an active patent is allowed to expire at alternative renewal dates conditional on having been previously traded; and (4) the probability that an active patent is allowed to expire at alternative renewal dates conditional on having been previously untraded. The joint set of moments describes both the history of the trading and the renewal decision of patent owners over the life cycle of patents.

The model evaluated at the parameter estimates fits the empirical moments reasonably well. The model captures both the probability that a patent is traded is decreasing with age and the previously traded patents being more likely to be traded than the previously untraded ones. The model also reflects that previously traded patents are less likely to be allowed to expire than the untraded ones. Fit is also assessed by the mean-squared difference between the empirical and the simulated moments, as well as figures that outline both the empirical and simulated moments pooled across patent grant years. Both methods show that the model evaluated at the parameter estimates fits well the empirical moments.

Several results emerge from the estimation. First, the value of the volume of the trade of patents represents about 50% of the total value of all patents. The relative value of the volume of the trade of patents is obtained by multiplying the number of patents traded in our simulation times the average value at age one of a patent traded, and then divide by the total value of all patents. In our analysis, the value at age one of a traded patent ($159,466) is about three times the value at age one of a non-traded patent ($50,507) (all values are in 2003 US dollars), and about 23% of the patents are traded at least once over their life cycle. The value at age one of a patent was estimated at ($74,684). This finding suggests that the gains from trade in the market for innovation may be significant, but the large differences between the average value of the traded and untraded patents also indicate that there may be selection in the trade of patents.

Second, we find that the gains from trade account for about 10% percent of the value of the volume of the trade of patents. The gains from trade were obtained by comparing the value of a patent actually traded with the value that the same patent would have obtained had the option

4

to sell it to a potential buyer not been allowed for over the patent's life cycle. An advantage of structurally estimating the model is that the counterfactual of shutting down the option of selling the patent allows us to quantify the gains from trade while accounting for selection in the transfer of patents.

Third, we show that the distribution of the gains from trade is highly skewed. We find that about 50% percent of the patents had gains from trade below $3,416, and these patents accounted for only 3.7 percent of the total value of the gains from trade. At the same time, the top 10% of the patents with the highest gains from trade (gains from trade higher or equal than $30,970 per patent) accounted for about 70 percent of the total value of the gains from trade. Furthermore, we demonstrate that the top one percent of the patents represented about 25% of the total gains from trade. This finding is what we expect, since higher quality patents are more likely to be traded and the distribution of the value of patents is known to be skewed (Schankerman and Pakes, 1986; Pakes, 1986). The relative small proportion of patents with significant gains from trade may have important implications on the workings and institutional settings of the market for innovation.

Fourth, the remaining of the findings refers to the effect that the lowering the costs of technology transfer has on the market for innovation. In a counterfactual experiment, we dropped the estimated sunk cost of technology transfer (estimated at about $5,500) by fifty percent. As an outcome to the experiment, here are two results we want to highlight. First, we found that the proportion of traded patents increases by six percentage points (from 23.1% to 29.6%). Second, we determine that the gains from trade increased by about 10%, and that the additional gains from trade come from patents with returns above the median and below the ninetieth percentile of the distribution of initial returns at age one. This finding, in our opinion, shows that the potential benefits associated with a potential reduction of the costs of technology transfer are limited.

Taken together, our empirical findings indicate that the market for innovation generates significant gains from trade, and that only a small proportion of the traded patents accounts for a significant share of the gains. Moreover, the costs of technology transfer affect this process by reducing the proportion of patents traded and by creating a selection in the transfer of patents, but the benefits of reducing these costs from the current levels on the realized gains from trade appear to be limited. Our interpretation is that as long as small firms can appropriate part of the gains from trade, this market increases the incentives to innovate; but that the potential benefits in gains from patent trades could be generated from reducing costs of technology transfer may be limited.

This paper contributes to both the literature on the measurement of the benefits of market

for innovation and to the estimation of the value of patent protection. The empirical documentation of gains from trade in the markets for innovation has been hard. Previous related work has provided estimates on the volume of trade of technology. This empirical analysis has typically involved strategic alliances' data and industry case studies. Arora, Fosfuri and Gambardella (2001) estimate the volume of trade in technology as 10-15 percent of non-defense R&D spending, and that this volume has grown over the last two decades. More recently, Robbins (2009) calculates that technology licensing by U.S. multinationals abroad accounted for about $60 billion of their receipts in 2002.[4] Both approaches face the challenge that the selection in the transfer of technology, such as the possibility that "the best" technologies are more likely to be traded and licensed, might generate a high volume of trade of technology while the realized gains from trade may be much smaller. Another challenge is that the licensing of patents is generally not publicly recorded; only patent licenses of significant value and involving publicly traded corporations are publicly reported. In addition, the costs of transfer of technology are not accounted for in previous empirical analysis. These potential issues are likely to be important in the market for innovation, given that the distribution of the value of innovation has been previously found to be skewed.

Our work also contributes to a well-known literature estimating the value of holding patents (Schankerman and Pakes, 1986; Pakes, 1986; Putnam, 1996; Lanjouw, 1998; Schankerman, 1998; Deng, 2007; and Bessen, 2008). In this literature, scholars have previously used the schedule of renewal fees and renewal decisions on patents to identify the distribution of the value of holding patents at alternative ages. Our work is distinct in that we use data on the transfer of patents at alternative ages, together with the renewal fees and renewal decisions, in order to recover the distribution of the value of holding patents at alternatives ages. The new data on patent sales allow us to build on their work in two ways. First, we estimate how much of the value of holding a patent is due to the possibility of selling the patents to potential buyers. In other words, we can uncover, under some assumptions, an estimate of the gains from trade in the market for innovation. Second, the identification of the parameters of the distribution of the initial per period returns is not uniquely based on observables from patents presumably in the left tail of this distribution (the renewal decisions), but also on observables of what appears to be high value patents (the traded patents). To the best of our knowledge, this is the first attempt quantifying the gains from trade using patent data, as well as to estimate the value of holding a patent using information on the transfer of patents.

This paper is organized as follows. Section 2 presents the model. Section 3 describes the

---

[4]Robbins (2009) uses internal records of the Bureau of Economic Analysis International Services trade data for licensing fees and royalty revenues reported on corporate tax returns. This estimate involves patent licensing royalties involving subsidiaries of the same parent firm.

estimation strategy. Section 4 presents the data sources and the patterns of the transfer and renewal of patents. Section 5 presents the estimation results of the distribution of patent value. Section 6 presents the results concerning the volume of the market for innovation and the gains from trade. Section 7 studies the effects of decreasing the costs of technology transfer on the proportion of patents traded and the gains from trade. Section 8 concludes the paper. All proofs are included in the Appendix.

2 A Model of the Transfer of the Ownership of Patents

This section presents the model of the transfer and renewal of patents to be estimated. We first describe the decision making problem that a patent holder faces, and then present new theoretical results that indicate how observations on the transfer and renewal decisions can provide information on the gains from trade, costs of technology transfer, and the distribution of the value of holding a patent.

The starting point for our theory is Schankerman and Pakes (1986) and Pakes (1986). They examine the problem of a patent owner deciding in each period whether or not to pay a renewal fee, and thereby extend the life of a patent, in a context with heterogeneity in the economic value of inventions. Our contribution is adding the possibility of the arrival of opportunities for surplus-enhancing transfers, which may lead to alternative potential owners having a greater valuation of a patent than the current owner. In addition, we consider that to transfer a patent to a new owner involves a transaction cost, a cost of technology transfer.

Alternative owners could generate a greater valuation than the current owner for a number of reasons. The economics of innovation literature has associated this additional valuation with vertical specialization (Teece, 1986; Arora, Fosfuri, Gambardella, 2002), complementarities in assets (Teece, 1986 and Gans and Stern, 2003), and comparative advantages in marketing or manufacturing (Arora and Ceccagnoli, 2006). As for the costs of technology transfer, previous scholars have identified them with technology adoption, expropriation risk, and the use of intermediaries such as patent lawyers (Arrow, 1962; Teece, 1977; and Astebro, 2002).

Assuming that both the transfer and renewal of patents are based on rational decision making at the patent level, patent owners will only sell their patents whenever the price obtained is higher than the value of retaining the ownership until the next period. Since transaction costs to reallocate the patent to the potential buyer are also a factor, the patent will only be sold if the added valuation is greater than the costs of technology transfer. Similarly, patent owners would renew their patents if the value of keeping the patent until the next renewal date is higher than

the renewal fee to be paid. If the renewal fees are not paid, the patent expires. If the patent owner pays the renewal fees and keep the ownership of the patent, the patent holder would face a similar problem next period. Finally, if a patent reaches its maximum legal length of patent protection, the patent will expire next period.

Patent owners maximize the expected discounted value of per period returns and the price they would obtain from selling their patent. Patent owners know the sequence of renewal fees and the costs of technology transfer, but are uncertain about both the arrival of opportunities for surplus-enhancing transfer and the sequence of per period returns that would be earned from retaining the patent. In this context, patent owners will not necessarily sell their patent to a buyer with higher returns because there is a positive probability that they will match in a future period with another buyer with much higher returns. Similarly, patent owners will pay the renewal fees and thereby retain for an extra period the patent ownership not only because the per period returns from retaining the patent may be high, but also because there is a positive probability that a future potential buyer will have much higher returns than them. The same decision problem will be faced in future periods with the exception that the patent horizon is shorter. An implication of the shorter horizon is that the patent holder will have less time to meet a buyer with high returns, and the potential buyer will have less time to amortize the sunk costs of technology transfer.

More formally, let $V_a(x, y)$ be the expected discounted value of patent protection prior to the transfer and renewal decision from a patent of age a with per period return x if kept by current owner and y if sold to the best potential buyer. If the renewal fee is not paid, the patent expires and $V_a(x, y) = 0$. If the patent is not sold and the renewal fee paid, the owner earns the per period return x and keeps the patent until the next renewal date. If the patent is sold, the buyer earns per period return y and pays the renewal fee. For simplicity, we will assume that the current owner obtains all the surplus in the sale (results are similar if there is Nash bargaining).[5] The value of holding a patent prior to the renewal and selling decision is:

$$V_a(x, y) = \max\{0, V_a^K(x, y), V_a^S(x, y)\}$$

where the age of the patent is $a \in \{1, ..., L\}$, L is the maximum legal length of patent protection, and $V_a^K(x, y)$ and $V_a^S(x, y)$ are the values of keeping and selling the patent, respectively. In

---

[5] The results may differ if the bargaining process is not efficient. Inefficiency could arise due to asymmetric information regarding x or y. Asymmetric information is less likely to be source of concern in the transfer of patents than in their licensing. For instance, if there was significant asymmetric information about x, the current owner and potential buyer could write a licensing agreement with an option to buy the patent. The option of licensing should decrease the degree of asymmetric information as of the time the patent is sold. In any case, if there was asymmetric information, the decision whether or not a patent had been previously traded would also be a state variable. The analysis of that decision problem, while interesting, is beyond the scope of this paper.

particular, $V_a^K(x,y)$, is the sum of the per period patent return x and the option value of the patent minus the renewal fee c_a. $V_a^S(x,y)$ is the price the current owner would obtain from selling the patent, which given the previous assumptions, is equal to the sum of the per period return that the buyer would earn y and the option value of the patent minus the renewal fee and the costs of technology transfer. Therefore we have

$$\begin{aligned} V_a^K(x,y) &= x + \beta E[V_{a+1}(x',y')|\Lambda_a] - c_a \\ V_a^S(x,y) &= y + \beta E[V_{a+1}(x',y')|\Lambda_a] - c_a - \tau \end{aligned}$$

where $\beta \in (0,1)$ is a discount factor, c_a is the renewal fee at age a, τ are the costs of technology transfer, Λ_a is the information set of the patent owner prior to the transfer and renewal decision in the ath year of the patent, and $E[.|\Lambda_a]$ is an expectation operator over the sequence of per period returns from internal (x') and external (y') opportunities conditional on the set of information Λ_a.

The following assumptions simplify the description of the decision making problem of a patent holder. We have divided these assumptions into three groups. The first group of them allows us to complete the description of the option value of the patent.

Assumption 1 (A1) $Pr(x',y'|\Lambda_a) = F_a(x',y'|x,y,d,\theta)$, where $Pr(.|.)$ denotes a conditional probability, $d \in \{K,S,0\}$ is the decision to keep the patent, sell it, or let it expire, and θ is a vector of parameters.

Assumption 2 (A2) Patent holders and potential buyers know the sequence of renewal fees c_a that will be required to be paid in order to keep the patent active. Furthermore, the renewal fee schedule in every renewal date is increasing with the age of the patent.

The option value of a patent now depends only on the current per period return of the patent owner (x), the return of the potential buyer (y), and the decision to keep or sell the patent (d). Assumption A1 implies that the past history of per period returns does not affect the future returns. Assumption A2 is consistent with the fact that renewal fees are increasing with the age of the patent at alternatives renewal dates. Therefore, under the above assumptions, the option value of keeping and selling the patent becomes

$$E[V_{a+1}(x',y') \mid x,y,d] = \int \int V_{a+1}(x',y') \mathrm{d}F_a(x',y'|x,y,d)$$

The second group of assumptions allows us to simplify the decision problem of a patent

holder. We will first motivate these assumptions, and then group them into assumptions A3-A5. Assumption A3 considers that the return under the best outside opportunity, y', depends on the current internal return, x'. This is motivated by the fact that while per period returns may be firm specific, it is also likely that they will be patent specific. Assumption A4 considers that the internal return x' depends on the previous realized return (x if $d = K$ or y if $d = S$), but it does not depend on the alternative return that was not realized (y if $d = K$ and x if $d = S$). This assumption is consistent with a framework where the arrival of new innovations depend on the innovations that have been adopted but not of those that were not. In assumption A5, we consider as true that the process of internal returns is the same for all patents with the same period return independently of whether or not the patent had previously changed ownership. This premise allows us to simplify the decision problem of patent owner, excluding the possibility that, conditional on patent returns, future patent returns may depend on the past history of trading decisions. An implication of the previous assumptions is that the law of motion $F_a(x', y'|x, y, d)$ can be rewritten as the product of the transition of external and internal returns $F_a(x', y'|x, y, d) = P_a^y(y'|x')[1_{\{d=K\}}P_a^x(x'|x) + 1_{\{d=S\}}P_a^x(x'|y)]$. These assumptions are summarized into Assumptions A3-A5.

Assumption 3 (A3) Let $F_a(x', y'|x, y, d, w) = P_a^y(y'|x', x, y, d) * P_a^x(x'|x, y, d)$ where P_a^y and P_a^x are probability functions. The return under the best outside opportunity, y', depends on the current internal return, x', but not on the previous history of returns or decisions: $P_a^y(y'|x', x, y, d) = P_a^y(y'|x')$.

Assumption 4 (A4) The internal return x' depends on the previous period realized return (x if $d = K$ or y if $d = S$) but not on the alternative return that was not realized (y if $d = K$ or x if $d = S$). That is:

$$P_a^x(x'|x, y, d) = \left\{ \begin{array}{ll} P_a^{x,K}(x'|x) & \text{if} \quad d = K \\ P_a^{x,S}(x'|y) & \text{if} \quad d = S \end{array} \right\}$$

Assumption 5 (A5) Conditional on the realized return at previous period, the internal return does not depend on the decision of keeping or selling the patent. That is:

$$P_a^{x,K}(x'|x) = P_a^x(x'|x)$$
$$P_a^{x,S}(x'|y) = P_a^x(x'|y)$$

Finally, the remaining two assumptions are made for convenience and simplicity, and allow us

to characterize the solution of the decision making problem of the patent holder. Assumption A6 considers that the costs of technology transfers are sunk. Furthermore, we suppose that they are independent of the per period patent returns and age. This assumption is motivated by the fact that previous scholars have identified these costs with large fixed noncapital and capital costs of adoption, organizational changes, and intermediaries such as patent lawyers and brokers (Teece, 1977 and Åstebro, 2002).

Assumption A7 considers that the transition probability of external and internal returns come from random walk processes. In particular, we have that the transition probability of internal returns, $P_a^x(x'|x)$ (if $d = K$) or $P_a^x(x'|y)$ (if $d = S$), comes from a random walk process $x' = g_a^i x$ (if $d = K$) and $x' = g_a^i y$ (if $d = S$) where g_a^i is a random variable non-increasing with a and independent of x. The assumption that the process of internal returns g_a^i is independent with x is consistent with the relationship between firm growth and firm size (firm growth is independent of firm size). We also need that patent returns do not increase too fast. A sufficient condition is that, for a given x, the probability that next period returns are higher than a given number is decreasing with patent age (A7.1). To parallel the characteristics of the process of internal returns, we assume that the process of external returns $P_a^y(y'|x')$ is also a random walk process such that $y' = g_a^{e'} x'$, where $g_a^{e'}$ is a random variable non-increasing with a and independent of x' (A7.2). We think of the improvement factor g_a^e as the arrival of opportunities for surplus-enhancing via the sale of patents relative to the in-house use and the licensing proceeds of the current owner. This is because the per period patent returns of the current owner may represent the proceeds both from using the technology in-house as well as licensing it to others.[6] That $g_a^{e'}$ is independent of the unobservable patent return x will allow us to prove that there exists a selling cutoff rule and to characterize its properties as a function of the patent age a and the patent return x.[7]

Assumption 6 (A6) Patent holders and potential buyer know the costs of technology transfer τ that will be required to be paid in order to transfer and adopt the patent. Furthermore, the costs of technology transfer are sunk and constant with the age of the patent.

---

[6]There are no systematic data on patent licensing revenue, but there is anecdotal evidence. IBM's licensing revenue was \$1.6 billion in the year 2000 (Berman (2002) as reported in Merrill, Levin and Myers (2004)). In 1996, U.S. corporations received \$66 billion in income from royalties of unaffiliated entities (Degnan, 1998). Texas Instruments reported to have obtained \$1.6 billion in licensing royalties from 1996 to 2003 (Grindley and Teece, 1997).

[7]The economic process we have considered here assumes that buyers of patents are adopters and users of the acquired technology rather than firms specializing in managing patents. For instance, a firm could exclusively focus on managing a patent by licensing it to many others who can then adopt it, but firms exclusively managing patents are a rather new organizational form mainly associated with firms acquiring patents for prospective litigation purposes.

Assumption 7 (A7) The transition probabilities $P_a^y(y'|x')$ and $P_a^x(x'|x)$ come from random walks processes. That is:

$$y' = g_a^{e'} x'$$

where $g^{e'}$ is a random variable non-increasing in a and independent of x' with cdf F^{g^e} that is continuous in y' at every x' such that, for a given x' and return z, the $P_a^y(y' > z|x')$ decreases with a. Similarly:

$$x' = \left\{ \begin{array}{ll} g_a^{i'} x & \text{if} \quad d = K \\ g_a^{i'} y & \text{if} \quad d = S \end{array} \right\}$$

where $g^{i'}$ is a random variable non-increasing in a independent of x with cdf F^{g^i} that is continuous in x' at every x such that, for a given x and z, the $P_a^x(x' > z|x)$ decreases with a.

2.1 The selection and horizon effect

To characterize when a patent holder will decide to sell, to keep, or lto et its patent expire, we must describe the properties of the value function and option value of a patent. The following Lemma shows that the value function of a patent is continuous, weakly increasing in the returns of the patent, weakly increasing in the return of potential buyers, and weakly decreasing in patent age.

Lemma 1. *The value function $V_a(x, y)$ is continuous and weakly increasing in the current return of the holder of the patent, x, and the return factor of the potential buyer, y. The option value $EV_{a+1}(x', y'|x, y, d)$ is weakly decreasing in a.*

Proof. See appendix. ∎

Based on the results from this Lemma, the following proposition shows that the solution to the problem of the firm can be summarized into two cutoff rules – $\widehat{x}_a(w)$ and $\widehat{g}_a^e(x, \theta)$ – that divide the state space into three areas (keep, sell or do not renew).

Proposition 1. *For each age, there exist cutoff rules $\widehat{x}_a(w)$ and $\widehat{g}_a^e(x, \theta)$ such that patents with an improvement factor above $\widehat{g}_a^e(x, \theta)$ will be sold. Among the patent holders that met a potential buyer with improvement factor below $\widehat{g}_a^e(x, \theta)$, the patents with per period returns x above $\widehat{x}_a(\theta)$ will be renewed, and those with returns x below $\widehat{x}_a(\theta)$ will be retained by their current owners.*

Proof. See Appendix. ∎
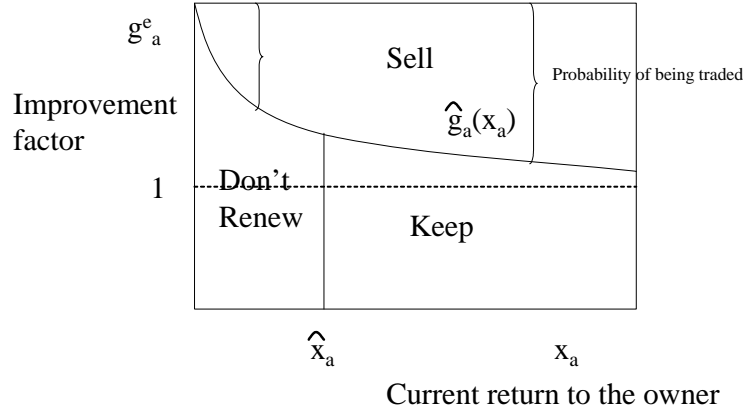
Figure 1: Optimal choices of a Patent Holder

Figure 1 presents the results of this proposition. In the vertical axis we have the improvement factor of the best potential buyer while in the horizontal axis we have the per period return of the current patent owner prior to the trading and renewal decision. The first cutoff is the per period return $\widehat{x}_a(\theta)$ that makes the owner indifferent between keeping the patent or letting it expire. This is a vertical line that separates the regions "Don't renew" and "Keep." This is the well known Pakes and Schankerman's renewal rule. The second cutoff $\widehat{g}_a^e(x, \theta)$ is new. This is the improvement factor that makes a patent owner indifferent between selling or not selling a patent with per period return x and age a. This is the curve that separates the region "Sell" from the regions "Don't renew" and "Keep." If a patent holder faces a potential buyer with an improvement factor above the cutoff $\widehat{g}_a^e(x, \theta)$, the patent will be sold. These patents belong to the "Sell" region. Moreover, as long as the improvement factor is lower than $\widehat{g}_a^e(x, \theta)$, patents with lower per period returns than $\widehat{x}_a(\theta)$ will be allowed to expire, this is the "Don't renew" region, and the ones with per period return higher than $\widehat{x}_a(\theta)$ will be renewed and retained by the current owner. This is the "Keep" region. The following proposition characterizes the properties of the function $\widehat{g}_a^e(x, \theta)$ and $\widehat{x}_a(\theta)$ with x and a.

Proposition 2. *The following properties characterize the functions $\widehat{x}_a(\theta)$ and $\widehat{g}_a^e(x, \theta)$:*

(a) *The function $\widehat{x}_a(\theta)$ is increasing with patent age a at the renewal dates.*

(b) *If the costs of technology transfer τ are positive, then*

(b.1) *For a fixed patent age a, the function $\widehat{g}_a^e(x, \theta)$ is decreasing with x;*

(b.2) *For a fixed patent return x, the function $\widehat{g}_a^e(x, \theta)$ is increasing with a.*

Proof. See Appendix. ∎

The findings in Proposition 2 are intuitive. In part (a), the renewal cutoff $\widehat{x}_a(\theta)$ increases with patent age a because while the fee schedule increases with a, the per period returns do not on average increase with a. This result is consistent with both Schankerman and Pakes (1986) and Pakes (1986). In part (b.1), we show that $\widehat{g}_a^e(x,\theta)$ is decreasing with patent returns (x). The new result is intuitive. If for a given improvement factor g^e the gains from trade by potential buyers are scaled up by the patent returns x, the higher is x the lower is the necessary improvement factor to amortize the costs of technology transfer. Lastly, part (b.2) of the proposition shows that for a fixed patent return x, the function $\widehat{g}_a^e(x,\theta)$ is monotonically increasing with patent age a. This is because for a given patent return x a shorter horizon implies less time to amortize the costs of technology transfer.[8]

Clearly, the functions $\widehat{g}_a^e(x,\theta)$ and $\widehat{x}_a(\theta)$, together with the distribution of improvement factor and the distribution of the per period returns of a patent at age a and prior the decision of the patent holder, determine the probability that an active patent will be sold, kept, or let to expire. The probability that a patent with return x will be sold at age a is the likelihood that an improvement factor is above $\widehat{g}_a^e(x,\theta)$.

An immediate implication from Proposition 2 is that there is both a selection and horizon effect in the trading of patents. There is a *selection effect* because the probability than an active patent is traded at age a increases with the patent return x. Note that as x increases the cutoff rule $\widehat{g}_a^e(x,\theta)$ decreases and thus the probability that a patent is being traded increases. The selection effect implies that traded patents, and especially those that have been recently traded, have on average higher returns than the previously untraded ones as of the time of the trading decision. This result is consistent with evidence on the transfer of patents. Patents with higher number of patent citations received are more likely to be traded (Serrano, 2010).

There is a *horizon effect* because, for a fixed patent return x, the probability that an active patent is being traded decreases with patent age. Note that as the patent age increases the cutoff rule $\widehat{g}_a^e(x,\theta)$ decreases with age and thus the probability that a patent is being traded decreases with age too because the random variable g_a^e is independent of x. The horizon effect accounts for why the proportion of active patents being traded may decrease with patent age. This result is also consistent with evidence on the transfer of patents.

---

[8]The key elements that allows Serrano to prove these results is that the costs of technology transfer is not proportional the to the difference between the value of selling and keeping a patnet. To the extent that a part of this costs is sunk, and thus does not fully internalize how the difference between the value of selling or not selling a patent changes as the revenue and age of the patent varies the result will hold. In other words, a part of the costs of technology transfer could principle depend on the current patent return x and the age of the patent a. On the other hand, if the costs of technology transfer was proportional to the difference between the value of selling and the value of keeping the patent, then neither part (b.1) nor (b.2) would hold.

2.2 Discussion of the results

The previous results suggest that observations on the transfer and renewal decisions can provide information on the distribution of the value of holding a patent, gains from trade, and the costs of technology transfer. First, observations of the proportion of patents allowed to expire and their corresponding renewal fees provide information to identify the distribution of patent value at alternatives ages. An implication of Proposition 1 and 2 is that the probability that an active patent is allowed to expire at age a is the proportion of active patents at age a with per period returns x above $\widehat{x}_a(\theta)$, and that the cutoff $\widehat{x}_a(\theta)$ is increasing with patent age at alternative renewal dates. Previous scholars have used the observed proportion of patents renewed at alternatives dates, together with the schedule of renewal fees, to estimate the distribution of the value of holding patents at alternatives ages.

Second, the selection into the trading decision suggests that observations on the trading of patents can provide additional information about the distribution of patent returns. At alternative ages we observe whether patents are traded, kept by their owners, or let to expire. Based on the selection effect, traded patents will have on average the highest returns of all. Patents retained by their owners will have, on average, lower returns than the traded patents. This is because patent owners holding patents with high per period returns will pay the renewal fees at alternative renewal dates not only because their current returns are high, but also because they also have a higher probability than the rest of patent holders to find a buyer with much higher returns. The rest of the patents, which will be let to expire, will have the lowest returns. Therefore, observations on the trading decision are also informative about the distribution of patent returns.

Third, observations of patents being traded can provide information about the resulting gainss. For a given distribution of per period returns at age a, and prior to the transfer and renewal decision, the observed proportion of patents being traded at age a will be associated with the proportion of buyers with improvement factors above the selling cutoff $\widehat{g}_a^e(x, \theta)$. Assuming that the transfer of a patent is based on decision making at the patent level characterized by the function $\widehat{g}_a^e(x, \theta)$, conditional on patent age and the distribution of patent returns, higher proportions of patents being traded will be associated with higher gains from patent trading. Thus, the observed proportion of traded patents at alternatives ages will provide some information on the distribution of the process of external returns (improvement factor), which determines the gains from trade in patents.

Fourth, observations of patents being traded at alternative ages can provide information about the costs of technology transfer. Conditional on the distribution of per period returns at alterna-

tive ages and the properties of the process of external returns, the decreasing pattern with age of the observed proportion of active patents being traded will be associated with the costs of technology transfer. Proposition 2 suggests that for a fixed patent return and a distribution of the process of external returns, the decreasing pattern of the observed proportion of patents being traded with age depends on the costs of technology transfer. Higher costs of technology transfer imply steeper decreasing patterns with age of the proportion of patents being traded.

Finally, the renewal rates of previously traded and untraded patents can provide additional information about both the gains from trade and costs of technology transfer. Conditional on the distribution of patent returns at alternative ages, the difference between the renewal rates of previously traded and untraded patents allows us to infer something about both the gains from trade of previously traded patents and the costs of technology transfer. In particular, higher costs of technology transfer and higher improvement factors, will imply a larger difference between the renewal rates of previously traded and the untraded patents. On the other hand, while higher costs of technology transfer imply that the proportion of patents being traded at a given age decreases; higher improvement factors will increase, on average, the proportion of patents being traded at a given age. Thus, information about the renewal rates of previously traded and untraded patents can provide information on both the process that determines the gains from trade and the costs of technology transfer.

2.3   Model stochastic specification

Initial per period patent returns   Following the literature on the estimation of patent value we assume that initial patent returns are lognormally distributed:

$$\log(x_1) \sim F_{x_1}(\mu, \sigma_R)$$

where $F_{x_1}(.)$ is a normal distribution. Previous scholars using survey data have shown that lognormal distributions provide a reasonable fit to the distribution patent returns (Harhoff, Narin, Scherer, and Vopel, 1999).

Transition of internal patent returns   The transition of internal returns is defined as

$$x_{a+1} = \left\{ \begin{array}{ll} 0 & \text{with probability } \gamma^i \\ x_a\delta & \text{if } d = K \\ y_a\delta & \text{if } d = S \end{array} \quad \text{with probability } (1 - \gamma^i), \right\}$$

where $d = K$ if the patent was retained by its current owner, $d = S$ if the patent was sold, and $\delta < 1$.

This stochastic process has an economic interpretation. At each age, patent owners and potential buyers find out whether patent returns may become obsolete ($g_a^i = 0$) and thus they become zero or alternatively the returns depreciate at a rate $\delta < 1$ ($g_a^i = \delta$). An advantage of this specification is that patent returns evolve stochastically. Conservatively, we have chosen to assume that the transition of internal patent returns is independent of patent age. Although this assumption undoubtedly represents a strong restriction on the process of internal returns, identifying age dependence in such stochastic process with U.S. patent renewal data may be challenging. This is because there are no renewal fees due until the fourth year following a patent grant date.[9]

Transition of external patent returns and costs of technology transfer We assume that there is a positive probability γ^e that the patent return earned by the potential buyer following a patent sale will be $y_a = 0$. Alternatively, there is a probability $(1 - \gamma^e)$ that the patent return earned by the potential buyer will be based on the improvement factor g^{*e} observed prior to the decision of selling the patent. In this case we will have that $y_a = g^{*e} x_a$, where the random variable g^{*e} is distributed exponentially with cdf $F^{g^{*e}}(g^{*e}) = 1 - \exp(-\frac{g^{*e}}{\sigma_a^e})$ and $\sigma_a^e = \sigma^e$. Consequently, the random variable g^e is distributed:

$$F^{g^e}(u) = \left\{ \begin{array}{cc} \gamma^e & \text{if} \quad u = 0 \\ (1 - \gamma^e)[1 - \exp(\frac{u}{\sigma^e})] & \text{if} \quad u > 0 \end{array} \right\}$$

Previous literature on the markets for innovation provides little guidance concerning a process to determine the gains from trading patents. An economic interpretation of the improvement factor g_a^e being independent of age can be that the exit of potential buyers that become patent owners is compensated by an arrival process of potential buyers (perhaps new sequential innovators) such that the distribution of the improvement factor of the best potential buyer is independent of the age of the patent. As for explicitly allowing for uncertain patent returns associated with the transfer of patents, previous scholars have reported that there are risks in the adoption of technology.[10]

---

[9] A more flexible stochastic process for g^i allowing for age dependence can, in principle, be estimated using both the renewal and the transfer data. Serrano (2006) argued that a stochastic process where g^i depended on age could be identified with information of the observed trading decisions prior to the first renewal date. In addition, Serrano (2006) used the flow of patent citations received by a patent at alternatives ages to estimate the value of holding a patent.

[10] The existence of uncertainty can also be thought of as re-interpretation of the benchmark model we previously analyzed. The costs of technology transfer could now be explained as having both a fixed (τ) and a variable

Measurement error   Because there may be measurement error in the transactions recorded at the USPTO, our specification allows for it in the transfer of patents. The specification we consider is the following. Let d^* and d be the true and the observed decision variables, respectively. If $d = 0$, then $d^* = 0$; If $d = K$, then $d^* = K$; but if $d = S$, then there is a probability $\varepsilon > 0$ that $d^* \neq S$.

3   Estimation

The parameters of the model are characterized by the vector $\theta = (\mu, \sigma_R, \sigma^e, \gamma^e, \tau, \varepsilon, \gamma^i, \delta)$. The discount factor is set to $\beta = 0.9$ as in Pakes (1986). The rest of the parameters are jointly estimated using the simulated generalized method of moments. This method involves finding θ in order to minimize 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 for analytically due to the fact that patent returns are unobserved and serially correlated over time.

3.1   Estimation algorithm

More formally, let the k parameter vector θ_0 be the unique solution to $G(\theta) = E[h - \eta(\theta)] = 0$, where the vector h is defined as the true hazard probabilities, and the vector $\eta(\theta)$ contains the probabilities predicted from the model given a parameter vector θ. Both h and $\eta(\theta)$ have dimension m, which is the number of moment conditions. Let N be the sample size. The sample moment condition is $G_N(\theta) = h_N - \eta_N(\theta)$. The vector $\eta_N(\theta)$ contains the hazards simulated by the model given the vector of parameters θ; and h_N represents the sample hazard proportions in the data.

The simulated minimum distance estimator, $\widehat{\theta}_N$, of the true parameter vector θ_0 is defined as

$$\widehat{\theta}_N = \arg \min_{\theta} [G_N(\theta, Z_N)]' W_N [G_N(\theta, Z_N)]$$

where W_N is a positive semi definite weighting matrix with dimension m by m defined as $W_N(\theta) = diag(\frac{1}{h_j^2} \sqrt{n_j/N})$, where n_j is the number of patents in which hazard j is conditioning on the relevant age/cohort, and h_j is the jth component of the vector of the sample hazard proportions.[11]

component $(\gamma^e y + \gamma^e \beta E[V_a(x', y')|x, y])$.

[11] The optimal weighting matrix would involve using the inverse of the asymptotic variance-covariance matrix of the sample moment conditions. Since efficient GMM is computationally very costly in our framework, we follow Lanjouw (1998) and use a diagonal matrix that weights each moment according to the number of observations of each sample hazard. Because the absolute value of the sample renewal rates approximately are ten times larger than the transfer rates, our weighting matrix divide each element in the diagonal matrix by its corresponding squared sample hazard. Thus, our minimization criterion is equivalent to one that minimizes the relative differences between

Given that the conditions required to have the consistency and asymptotic normality of our estimator are satisfied (Pakes and Pollard, 1989), and that $W_N(\theta)$ converges in probability to a semi-definite matrix W_0, we have that $G_N(\theta)$ converges in distribution to a normal distribution,

$$\sqrt{N} G_N(\theta_0, Z_N) \to N(0, V)$$

and $\widehat{\theta}_N$ converges to θ_0, and $\sqrt{N}(\widehat{\theta}_N - \theta_0)$ satisfies a central limit theorem

$$\sqrt{N}(\widehat{\theta}_N - \theta_0) \to N(0, (\Gamma' W_0 \Gamma)^{-1} \Gamma' W_0 \Omega W_0 \Gamma (\Gamma' W_0 \Gamma)^{-1})$$

where

$$V = (\Gamma' W_0 \Gamma)^{-1} \Gamma' W_0 \Omega W_0 \Gamma (\Gamma' W_0 \Gamma)^{-1}$$

$$\Gamma = p \lim \frac{\partial G_N(\theta_0, Z_N)}{\partial \theta'}$$

$$\Omega = E([G_N(\theta_0, Z_N)][G_N(\theta_0, Z_N)]')$$

3.2 Obtaining $\widehat{\theta}$

There are a number of steps we follow to obtain the estimate $\widehat{\theta}$. First, we simulate the model in order to obtain the vector of hazard probabilities $\eta_N(\theta)$. This involves obtaining the cutoff rules $\widehat{x}_a$ and $\widehat{g}_a^e$ and the value function of keeping and selling the patent. The second step involves finding the value of θ that minimizes a metric of the distance between the simulated and the empirical hazard probabilities.

Given a vector of parameters θ, the hazard probabilities $\eta_N(\theta)$ are generated by first solving the model and then simulating it. To solve the model, we compute both the value of keeping a patent and the cutoff rules of renewing and selling it. With the purpose of doing so, we solve the model recursively, starting from the maximum legal length of patent protection and moving backwards until the patent's grant date. To obtain the value function of keeping the patent at alternative ages, we discretize the per period return into a number of grid points.[12] Solving the value function backwards involves constructing cutoff rules for selling and renewing a patent. At each patent age we use this cutoff rules in order to approximate the integral that defines the

---

the sample and the simulated hazards.

[12] We apply a grid of 10 points, but have used 100 and 300 grid points and found no significant differences. Rather than the number of points used in the grid, the important aspect is to chose their location based on the curvature of the value function. The location of the grid points was chosen to match a lognormal distribution with parameter similar to those of the current estimates of initial patent per period returns. This allowed us to ensure that the density of the grid points is higher where the curvature of the value function is steeper.

expected value of patent protection (i.e., the patent's option value).[13] The backward process ends when we reach the patent's age one. At this point, we have calculated the value of keeping a patent at alternative ages and the grid points of the per period patent returns.

The life cycle history of N patents is simulated S times using the previously calculated values of keeping and selling the patent. Simulating the life cycle of these patents consists of taking pseudo-random draws from the distribution of the initial returns, and then pass each initial patent return through the stochastic processes of internal and external patent returns, including the measurement error in recorded transfers. History of trade and renewals decisions for each patent, as well as their patent returns and patent value, are obtained and recorded. The simulated hazard probabilities $\eta_N(\theta)$ are calculated by averaging out, over the number of patents and simulations, the patent renewal and transfer decisions.

Finally, an algorithm based on simulated annealing methods is used to obtain the simulated minimum distance estimate. Annealing methods are used because the objective function is likely to be non-smooth and have multiple local minima.[14] Parameter estimates and their respective standard errors were obtained by applying a Newton method algorithm once the simulated annealing method converged to the global minimum of the objective function of the estimator.

4 Data

Our empirical analysis uses data we obtained from the United States Patent and Trademark Office (USPTO) for the period 1981-2002. The key variables for our study are the indicators of two important events in the lifetime of a patent: renewals and assignments.

Patent renewal data Patent applications become patents as of their grant date.[15] Multiple patent renewal fees must be paid to extend the life of a patent until its maximum legal length. If a renewal fee is not paid at a renewal date, then the patent immediately expires.[16] A number of scholars have used the drop out rates at alternative renewal dates to estimate the distribution of the value of holding patents (Schankerman and Pakes, 1986; Pakes, 1986, and others). In the US, patents are subject to renewal fees only if they were applied for after December 12, 1980. We have obtained these records, as of December 31, 2001, for patents applied for since December 12,

---

[13]We use quadrature methods to numerically approximate the value of the integral.

[14]In particular, we use the simulated annealing algorithm developed and tested in Goffe, Ferrier and Rogers (1994).

[15]The maximum legal length of new patents applied for prior to 1995 was 17 years following their grant date. This maximum legal lenght was subsequently modified to 20 years following patent application date. During ther period 1981-2002 patent applications were granted on average in 2.5 years.

[16]The USPTO began charging renewal fees in 1984 on patents applied for after December 12, 1980.

1980 and granted after 1983 until the end of 2001. Similarly to the European patent data, the amount of renewal fees is increasing with the age of the patent.[17] However, whereas renewal fees in Europe are characterized by yearly renewal dates and having all patentees paying the same fee, renewal fees in the US are characterized by three renewal dates (at the end of 4th, 8th and 12th years following a patent's grant date) and having patentees paying different fees depending on whether their entity status is small or large. The USPTO defines small entities as individual inventors, non-profit organizations, universities, government agencies, and small businesses (firms with less than 500 employees). Large entities are the rest of patentees, primarily large businesses (firms with more than 500 employees). Small entities are subject to renewal fees that are half of the amount than large entities pay.

Patent assignment data   Another event that can happen during the life cycle of a patent is what the USPTO calls "assignments." We follow Serrano (2010) and use re-assignment data to identify transfers of patents across firm boundaries. The source of this data is the USPTO Patent Assignments Database. When a U.S. patent is transferred, an assignment is recorded at the USPTO acknowledging the change in ownership. We have also obtained these records for all transfers that occurred from 1981 to 2002. 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.[18] The procedure conservatively drops all the assignments that appear not to be associated with an actual patent trade. The details of the procedures we used to deal with the assignment data are explained in Serrano (2010). Most relevant for this analysis is that the remaining assignment records identify the sale of a patent and have information about patent numbers, making it possible to merge them at the patent level with the payment of renewal fees, the patentee's entity status, and other patent characteristics.[19]

Under Section 261 of the U.S. Patent Act, recording the assignment protects the patent owner against previous unrecorded interests and subsequent assignments. If the patentee does not record the assignment, subsequent recorded assignments will take priority. For these reasons, patent owners have strong incentives to record assignments and patent attorneys strongly recommend this practice (Dykeman and Kopko, 2004).

---

[17]Most of the previous literature used European patent data (starting with Schankerman and Pakes, 1986 and Pakes, 1986).

[18]We also dropped records in which the buyer and seller are the same entity and in which the execution date is either before the application date or after patent expiration. For additional details on the procedure, see Serrano (2010).

[19]Lamoreaux and Sokoloff (1997, 1999) and this paper are the only ones that explore patent assignments to study markets for innovation. They used a sample of sales of private inventors patents and provided a historical account of whether organized markets for technology existed in the late 19th and early 20th century.

The dataset we have compiled is a panel of patents applied for after December 12, 1980 and issued to US patentees since January 1, 1983 that details their history of trade and renewal decisions until 2001.[20] The data contains patent characteristics such as their grant year, the patentee's entity status, etc.

Our empirical analysis will focus on small firms. We operationalize this focus by restricting attention to patents granted to firms and applied for businesses with no more than 500 employees. In doing so, the economic forces that we highlight will be more salient than in transactions involving the sales of patents by large corporations. About 78 percent of all the patents in the period of analysis were granted to firms; the rest were granted to government agencies, individual inventors, or were unassigned as of their issue date. In our dataset, small businesses patents account for approximately 15 percent of all the patents granted to firms. They are more likely to be traded and to receive patent citations than the large business patents. In addition, small firms are interesting in their own right, given the important role they play in the innovation process (Arrow, 1983 and Acs and Audretsch, 1998).[21]

Table 1 presents a summary of the renewal and the transfer rates of patents in our dataset. The top part of the table presents the expiration rates of active patents at alternative renewal dates for previously traded and untraded patents. Expiration rates of previously traded patents are lower than the expiration rates of the previously untraded patents. In addition, the table shows that the expiration rates increase with patent age. The bottom part of the table presents the probability that an active patent is traded at alternatives ages for both the previously traded and the untraded patents. The probability of trade of previously traded patents is higher than the probability of trade for patents that were previously untraded. Furthermore, the table shows that the probability of trade, whether or not the patent was previously traded, decreases with age. In our previous work we showed that these patterns were consistent across type of patentees and patent technology classes.[22]

---

[20]The advantage of focusing on U.S. businesses is that the firms are more comparable. The disadvantage is that the sample of patents decreases significantly and that there is a decreasing time trend in the patenting of U.S. businesses at the USPTO.

[21]Alternatively, we could have considered small innovators as in Serrano (2006, 2010). A small innovators patent is patent owned by firms that were granted no more than five patents in a given year. Small innovators and small business patents are highly correlated. The advantage of focusing on small businesses, instead of small innovators, is that we categorize firms based on a traditional measure of firm size. In addition, the magnitude of patent fees to be paid, including patent renewal fees are based on whether or not the firm is small or large businesses, and not on the number of patents obtained. The disadvantage of using small businesses is that it likely that some of them are highly innovative (i.e., patent often) while others may not. The number of patents granted and applied for small businesses is about half of the patents granted to small innovators, and small innovators represented about thirty-three percent of all the patents granted to firms.

[22]Interestingly, the findings are similar independently of whether firm size is measured by patent counts (as in Serrano, 2010) or number of employees.

Table 1: Percentage of Active Patents Traded and Expired

Age	All	Not Previously Traded	Previously Traded (Years since last trade)	
			Any Year	One year
A. Probability that an active patent is allowed to expire				
5	17.3	17.9	12.9	6.8
9	27.9	28.8	23.6	12.3
13	28.4	29.0	26.6	19.4
B. Probability that an active patents is traded				
2	3.16	2.88	7.21	7.21
7	2.73	2.31	4.52	6.37
11	2.41	2.13	3.40	2.74
16	2.36	1.13	2.98	-

Estimation sample Our estimation sample uses patents granted during the period 1988-1997. This is a sub-sample of our dataset. Focusing on these years allow us to minimize the effects of the multiple patent policy changes during the last few decades. The most important patent policy events were: (a) the drastic changes in patent fees during the early 1980s by US Congress; (b) the perception that by the year 1986 the establishment of the Court of Appeals for the Federal Circuit (CAFC) reflected a "pro-patent" shift; and (c) the Patent Law Act of 1996 redefining and extending the maximum legal length of patent protection.[23] These policy changes likely affected the characteristics of the patents being applied for and thus the distribution of the initial returns to patent protection.[24] In the sample we use to estimate the model the distribution of patent values will be less affected by these policy changes.

There are 54, 840 patents distributed across ten cohorts in our estimation sample. Each cohort is defined as the grant year of a patent. There are 20, 790 patents granted for the years 1988-

---

[23] A challenge in using time series of renewal data to estimate a distribution of patent values is that changes in patent policy may affect the characteristics of innovations that firms applied for. During the last three decades there have been three major changes in U.S. patent policy. First, patent fees were subject to drastic changes by U.S. congress during the mid 1980s. We believe these changes affected the type of innovations being applied for and thus generated a non-stable distribution of the returns of patents applied for the period 1983-1987. Second, the Court of Appeals for the Federal Circuit (CAFC) was established by the U.S. Congress in 1982. The decisions of the CAFC were widely viewed as reflecting a "pro-patent" shift by the year 1986 (Hall and Ziedonis, 2001; Jaffe and Lerner, 2004). We also think that these changes likely affected the expected value of obtaining a patent, and thus the characteristics and the number of patents that firms applied for. This likely lead to changes in the distribution of initial per period patent returns. Finally, the Patent Law Act of 1996 extended the maximum legal length of patent protection for patents applied for after 1996. The maximum legal length was changed from a maximum of seventeen years following the patent's grant date, to a maximum of twenty years following the application date. This change may have affected not only the value of patents, but also the characteristics of innovations being applied for. Previous scholars have argued that the time from patent application to grant is not exogenous. It depends on the characteristics of the patent such as complexity of the technology, which may be correlated with patent value.

[24] While the analysis of these changes in patent policy may be interesting on its own, we have considered that accounting for these patent policy changes in our estimation strategy is outside of the scope of our study.

1992 and 34,050 for the rest of the cohorts. Each patent is observed on average 8.5 years in our sample, but the maximum and minimum amount of time the patents can be observed is thirteen and four years respectively depending on the patent's grant year. The number of years a patent is observed during its life cycle differs across cohorts because our transfer and renewal data ends in 2001. Patents granted in 1988 are observed from their issue date to age thirteen, and for the patents granted in 1997 we only have data for the first four years. For each of these granted patents we will observe the trading and renewal decisions at alternative ages over their life cycle.

In the estimation of the structural parameters of our model, we maximize a criterion function that measures the goodness of fit of the model to the following stylized facts:

1. The probability that an active patent is allowed to expire at a renewal date for previously traded patents;

2. The probability that an active patent is allowed to expire at a renewal date for previously untraded patents;

3. The probability that an active patent is being traded at a given age for previously traded patents;

4. The probability that an active patent is being traded at a given age for previously untraded patents.

The dimension of the vector of moment conditions will be $m = 186$, which is the sum of 36 conditional probabilities based on the renewal dates and 150 conditional probabilities based on the transfer dates.[25]

5   Estimation Results

Table 2 presents the parameter estimates and the corresponding standard errors. On the top part of this table, estimates for the initial distribution of patent returns as well as the processes of depreciation and obsolescence can be found. The bottom part shows the parameter estimates for the process of the gains from trade and the cost of technology transfer. The parameter estimates are all positive and significant.

---

[25]The 1988 patent cohort will count towards twelve probabilities in the trading decision (i.e., from age two to thirteen), and three in the renewal dates for each the previously traded and the untraded patents. The 1997 patent cohort, instead, will count towards three probabilities in the trading decision (i.e., from age 2 to four) and one in the renewal decision (i.e, the first renewal date).

An indicator of how the estimated model fits the data is comparing the empirical moments and the simulated moments from the model. Figures 2 and 3 (see appendix) outline both the empirical and simulated moments pooled across patent grant years. The model fits the empirical moments used in the estimation algorithm reasonably well. Figure 2 presents both the simulated and empirical probabilities that an active patent is allowed to expire for previously traded and untraded patents. The probability of an active patent is being allowed to expire conditional on having been previously untraded is well fitted by our model simulations. For the untraded patents, however, the simulation generated by the model tends to predict less steepness than what is observed in the data, especially in the second and last renewal date. This feature may be due to unobserved patent heterogeneity in that some patents are more likely to be traded than others independently of their patent returns. Alternatively, it could be due to our rather simple specification of the process of the transition of internal returns. For another moment that sheds light on how the estimated model fits the data, Figure 3 presents the probability of that an active patent is traded for previously traded and untraded patents.[26] The model is able to capture well the decreasing shape of both probabilities with age.[27] Another indicator of the fit of the model is the sum of squared residuals of the difference between the estimated and the actual moment conditions divided by the number of moment conditions (MSE). At the bottom of Table 2, we show that the MSE is low $(4.2447 * 10^{-4})$.[28] Putting these findings together, both indicators suggest that our estimated model fits the data well.

Parameters μ and σ_R determine the initial distribution of the per period returns as of the grant date of patents. A high σ_R implies a high degree of heterogeneity in initial patent returns. A low μ implies that the initial per period returns to patents are low. The estimates of μ and σ_R indicate that a large proportion of the patents in our data will start with very low per period returns. Parameters δ and γ^i represent depreciation and obsolescence with age of patent returns, respectively. The estimates of δ and γ^i imply that the per period returns of a patent decrease on average at the rate of approximately thirteen percent a year. These parameters together accounts for the distribution of initial returns and its evolution over the age of patents.

Table A-1 presents a summary of the distribution of per period returns at ages one, four, and seven. This table and the remaining ones are in the appendix. The descriptive statistics provided in this and the subsequent tables were obtained from a simulation run of $90,000$ patent draws

---

[26]Note that the age profiles we examine are right censored at age thirteen since this is the longest age profile of the oldest patent cohort in our estimation sample.

[27]In out of sample predictions, we also found that the probabilities sharply decrease when patents get closer to the maximum legal length of patent protection (age 17 for patents applied for before 1997), which is consistent with the evidence on the transfer of patents presented in section 2.

[28]This magnitude of the MSE is comparable to the reported MSE in Pakes (1986).

Table 2: Parameter Estimates

Description (Parameter)	Estimate[a]
A. Patent initial returns, depreciation, and obsolescence	
Depreciation factor (δ)	0.9003 ($4.35*10^{-2}$)
Full Depreciation (γ^i)	0.9702 ($7.60*10^{-3}$)
Mean parameter of the Lognormal Initial Distribution (μ)	8.3115 ($4.50*10^{-3}$)
Std. Deviation parameter of the Lognormal Initial Distribution (σ_R)	1.7202 ($4.30*10^{-3}$)
B. Market for technology and costs of technology transfer	
Cost of Technology transfer (τ)	5,494.8 (63.46)
Mean External Growth of Returns (σ^e)	0.3775 ($2.70*10^{-3}$)
Proportion of unsuccessful transfers (γ^e)	0.0386 ($2.86*10^{-4}$)
Random transfers (ε)	0.0053 ($8.15*10^{-4}$)
Size of sample	54,840
Simulations in the estimation	164,520
MSE[b]	$4.2447*10^{-4}$

[a]Estimated standard errors in parenthesis.

[b]MSE is the sum of squared residuals of the difference between the estimated and the actual moment conditions divided by 186, the number of moment conditions.

using the estimates in $\widehat{\theta}$ and the renewal fee schedule of the year 1988 (in 2003 dollars).[29] There are two results in this table that are worth highlighting. First, the left tail of the distribution of per period returns becomes thicker with age. Per period returns decreased and a significant proportion of patents became obsolete between ages one, four, and seven. Eleven percent of the patents at age one had per period returns below $500; this proportion doubled (22.6 percent) by age four, and was 33.2 at age seven. The median of the distribution of per period returns ($4,080$) at age one dropped by approximately half ($2,447$) by age four. The median per period return dropped to $1,404$ by age seven. The second result to note is that the distribution of per period returns at age one is highly skewed. At the same time that about twenty percent of the patents have returns at age one below or equal $1,000$, only the top 3.5 percent had returns above $100,000$. These findings suggests a highly skewed distribution of patent value.

In Table A-2, we provide the percentiles of the distribution of patent values at age one, four and seven. The value of a patent is defined as the discounted present value of the per period returns (minus the renewal fees if the patent is renewed, and minus the costs of technology transfer if the patent is sold) from age one, four or seven to the maximum legal length of patent protection. Column (1) presents the patent value as of age one. Twenty-five percent of the patents in a patent

---

[29]Alternatively we could have used the average of the renewal fee schedule across all the year in the sample. We do not expect the results to be different.

cohort had values below or equal to $4,584. These patents accounted for 0.7 percent of the total value of patents in a cohort, while the bottom fifty-percent of the patents contributed a slightly less than 4 percent of the total value in a cohort. The value of the median patent at age one was $16,184 and the mean $76,598. The top 10 percent of the highest valued patents accounted for approximately seventy percent of the total value of a cohort. These findings confirm a very skewed distribution of patent values. The rest of the columns in the table present the distribution of patent value at ages four and seven. A last result we want to highlight is that the distribution of patent value appears to become more skewed with patent age.

Parameters γ^e and σ^e jointly determine the potential benefits of transferring patents. Recall that there is a probability γ^e that after a patent trade the returns are zero. Alternatively, with probability $(1 - \gamma^e)$ the patent returns are based on the improvement factor observed prior to the transfer of the patent. The low estimate of γ^e (about 0.0386) indicates that in most of the transfers the realized returns depend on the observed improvement factor of the potential buyer. In other words, the uncertainty of the potential benefits of acquiring a patent involves about a four percent premium.[30] The improvement factor that determines the returns of a buyer is a function of the parameter σ^e. A high σ^e is associated with higher improvement factors and thus higher patent returns. The estimate of σ^e implies that with probability 0.071 the patent returns of the buyer in a given period are higher than the ones of the current owner. Furthermore, conditional on a buyer having higher returns than a current owner, the buyer's per period returns are thirty-eight percent higher than the ones of the current owner. These findings suggest important benefits of reallocating the ownership of a patent from the original inventor to a new owner for whom the patent has a higher value. Furthermore, the low estimate of ε implies that about half of a percentage point of the simulated proportion of active patents can be accounted for our specification of the measurement error in the trade of patents.[31]

Another important point to note is that the estimate of the costs of technology transfer τ is significant (about $5,500). The sunk costs of technology transfer are equivalent to a third of the estimated average patent return at age one. Alternatively, these costs represent about eight percent of the average present value of a patent at age one. The magnitude of the estimated costs of technology transfer points at the possibility of significant selection effects into the trading of patents.

---

[30]Given our model specification, an alternative interpretation could be that the variable component of the costs of technology transfer represent about a four percent of the sale price of a patent.

[31]An alternative is that half of a percentage point of the simulated proportion of the traded patents can be accounted for a transfer process that does not involve neither sunk nor variable costs of technology transfer.

6　The market for innovation and gains from trade

This section documents our empirical results for the volume of the market for innovation and the gains from trade. We first compare the value of the traded and untraded patents. We then estimate the volume of the market for innovation, which is the total value of the traded patents. We end the section by quantifying how much of this volume is gains from trading patents.

In Table A-3A, we present the distribution of patent value at age one for both the traded and the untraded patents. Traded patents are those that were sold at least once over their life cycle; untraded patents are the rest. The first result we want to highlight is that the mean of the distribution of the value of patents eventually traded ($164, 670$) is about three times larger than the mean of the distribution of value of the patents untraded ($50, 162$). The finding that traded patents have a higher value than the untraded ones is consistent across the percentiles of the distribution of the value of the traded and untraded patents. Furthermore, the difference between the value of the traded and the untraded patents increases over the percentiles of the distributions. Differences range from few thousand dollars at the fifth percentile of each distribution (654 and $4, 877$ for untraded and traded patents, respectively) to about two million dollars at the 99th percentile ($1, 687, 702$ and $3, 754, 001$).

Table A-3B quantifies the volume of the trade of patents. We consider two indicator variables for transactions in the market for innovation: the proportion of patents traded and the value of the traded patents. In the first row we report that the traded patents represent 23.1 percent of all the patents.[32] In other words, among all the patents in our simulation, twenty-three percent of them were traded at least once over their life cycle. The next row shows that the value of the traded patents represents 49.6 percent of the total value of patents as of age one. This estimate correspond to the total value of the patents traded as a percentage of the value of all the patents. This result indicates the importance of the market for innovation. The volume of the market for US small firm granted patents in between 1988-1997 is then approximately $2.08 billion[33]

In Table A-4, we begin tackling the issue of selection in the trade of patents. The gross differences between the value of the traded and the untraded patents (and thus the value of volume of the trade in patents) can be due to potential gains from trade, but also because patents with higher returns are more likely to be traded. Column (1) presents the probability that a patent is traded obtained from simulations based on the estimated costs of technology transfer.

---

[32] The number of traded patents is equal to the proportion of traded patents (0.23) times the number of patents in our estimation sample (54, 840).

[33] This is the number of patents (54, 840) × the proportion of patents traded (0.231) × the value at age one of a patent eventually traded ($164, 670$).

The bottom row shows the sample probability that a patent is traded over the life cycle, while the rest of the rows report the same probability across groups of patents formed based on the percentiles of the distribution of per period returns at age one. The sample probability that a patent is traded (0.231) is more than four times larger than the probability of patents with per period returns at age one in between the 10th and 25th percentile (0.05). Furthermore, patents with per period returns at age one below the 10th percentile are virtually not traded. In contrast, the trade of patents is much more common at the highest percentiles of the distribution of per period returns at age one. For instance, forty percent of the patents in between the 75th and 80th percentiles are traded over their life cycle. These results confirm that there is selection in the trade of patents and suggests that this selection can account for a sizable part of the value of the volume of the trade of patents.

Table A-6A quantifies how much of this volume of the trade in patents is due to gains from trade. The gains from trade were obtained by comparing the value of a patent actually traded with the value that the same patent would have obtained had the option to sell it to a potential buyer not been allowed for over the patent's life cycle. Column (1) and (2) report the average patent value at age one of traded patents with and without the possibility to transfer it to buyers. In Column (3) we present the average realized gain from trade as a percentage of the value of a traded patent across the percentiles of the distribution of patent returns at age one. Gains from trade for the average traded patent represent an additional 10 percent to the value that the original patentee could have earned by keeping the patent. We also show that the gains from trade as a percentage of patent value are heterogeneous across patents based on their initial returns at age one. The average gains from trade range from as low as zero for traded patents with per period returns below the 10th percentile to as high as 11 percent for patents with per period returns at around the 99th percentile. For traded patents with per period returns at age one in between the 50th and 75th percentile, about 9 percent of their value was due to gains from trade. An advantage of structurally estimating the parameters of the model is that a simple counterfactual allow us quantify the gains from patent trading while at the same time account for selection in the trade of patents.

In Table A-5, we assess in more detail the skewness and the distribution of the value of the gains from trade. This table provides percentiles and cumulative percent of the distribution of the gains from trade of the traded patents. The result we want to highlight here is that the distribution of the gains from reallocating patents to new owners is highly skewed. The table reports that fifty percent of the traded patents had gains from trade below or equal $3,416$ (the mean of the value of the trade of patents is about $15,008$). These patents accounted for only 3.7

29

percent of the total value of the gains from trade of patents in a cohort. On the other hand, the top ten percent of the patents with the largest gains from trade accounted for about 68 percent of the realized gains from trade. These findings indicate that only a small share of the patents contributes significantly to the realized gains from trade.

These findings suggest that the market for innovation generates significant gains from trade for small firms, but that only a small proportion of the traded patents account for significant share of the realized gains from trading patents.

7   Effects of costs of technology transfer on the trade of patents and gains from trading patents

The remaining of our analysis studies the effect of lowering the costs of technology transfer on both the proportion of patents traded and the gains from trading patents. To this extent, we decrease the costs of technology transfer by fifty percent, i.e., from \$5,495 to \$2,747, while keeping the rest of the parameters at their original estimated levels.

Several results emerge from this counterfactual experiment. First, we study the effects of reducing the costs of technology transfer on the probability that a patent is traded. Column (2) in Table A-4 reports the probability that a patent is traded across groups of patents based on their initial returns at age one, as well as the overall sample probability that a patent is traded. First, there are two aspects worth highlighting here. The probability that a patent is traded increases by six percentage points (from 0.231 to 0.296). Moreover, we show that the change in the probability that a patent is traded is heterogeneous across patents based on their initial returns at age one. Patents with returns in between the percentiles around the median of the distribution of initial returns at age one (25th and 75th) increased their probability of trade by about ten percentage points, while the rest of the patents were much less affected by the lower costs of technology transfer.

Second, in Table A-7, we present the effects of lowering the costs of technology transfer on the gains from trading patents. Column (1) presents the percentage change in the value of the gains from trade for all patents and across groups of patents based on the initial returns at age one. In Column (2), we present the cumulative distribution of the additional gains from trade of the traded patents across groups of patents based on their initial returns at age one, and in Column (3) we show the cumulative distribution of the gains from trade prior to the reduction of the costs of technology transfer across the same groups of patents. There are two findings we especially want to highlight. First, in Column (1) we show that the gains from trading patents

increase by about 10%. That is, decreasing the costs of technology transfer by fifty percent have limited increases in the total value of the gains from trade. Second, in Column (2) we show that the source of the additional gains from trade are patents with returns in between the median and the seventy-five percentile of the distribution of initial returns at age one. While the bottom seventy five percent of the traded patents with the lowest returns at age one account for 38% of the additional gains from trade, the patents with returns at age one in between the fifty and seventy percentile account for about 24% to the additional gains from trade. In contrast, in Column (3) we show that the patents with the lowest seventy-five percent of initial returns at age one accounted for only 6% of the gains from trade. These findings suggest that the potential benefits associated with reductions in the costs of technology transfer are limited, but that the patents that will benefit the most are those with returns at age one somewhat above the median of the distribution of initial returns at age one.

Taken together, the findings of this counterfactual indicate that while the proportion of patents traded increases significantly, the benefits of reducing the costs of technology transfer on the realized gains from trading patents appear to be limited.

8   Conclusion

This paper develops and estimates a model of the transfer and renewal that, under some assumptions, allows us to quantify the gains from trading patents and the costs of technology transfer in the market for innovation. The model is estimated using new data on the transfer of patents in a sample of patents applied for and granted to US small firms.

Several findings emerge from our analysis. First, the volume of the trade of patents represents about fifty percent of the total value of all patents. Second, the gains from the trade of these patents represent about ten percent of the volume of the trade of patents. Third, we showed that the distribution of the gains from trade was highly skewed, and that a small fraction of the traded patents accounted for a large part of the realized gains from trade. Fourth, we studied the effects of lowering the costs of technology transfer by fifty percent, and found that the probability that a patent is traded increases by six percentage points and the value of the gains from trade rises by only ten percent.

The findings in this paper indicate that the market for innovation generates significant gains from trade for small firms, but that only a small proportion of the traded patents accounts for a large share of the benefits. Our interpretation is that as long as small firms can appropriate part of the gains from patent trading, this market increases the incentives to innovate. The result that

only a small proportion of patents will generate significant gains from trade may have important implications in the institutional settings and market structure of the intermediaries of the market for innovation. Moreover, we found that the costs of technology transfer affect the workings of the market for innovation by reducing the proportion of patents sold and generating at the same time a selection in the trade of patents. There seem to be, however, limited benefits from reducing these costs from the current level.

At the same time, this paper may have some limitations. We likely examine a domain where the sale of patents may be considerably more useful than for the larger firms. In this respect our estimates may overestimate the overall gains from patent trading in the economy. Nevertheless, our model and empirical methodology is with ease portable to other related domains such as large firms. At the same time our sample does not contain the licensing of patents, and thus the benefits of the market for innovation could be much larger. A limitation of our setting is that it does not allow for the possibility that firms may choose to increase patenting as well as to specialize in what they excel. Our data also exhibited some limitations such as not linking the type and the name of firms acquiring patents, not reporting the sale prices, and not collecting the contractual terms in the patent transactions. These limitations provide opportunities for further research.

References

ACS, Z. J., AND D. B. AUDRETSCH (1988): "Innovation in Large and Small Firms: An Empirical Analysis," *American Economic Review*, 78 (4), 678–690.

ARORA, A. (1997): ""Patents, Licensing, And Market Structure In Chemicals," *Research Policy*, 26, 391–403.

ARORA, A., AND M. CECCAGNOLI (2006): "Patent Protection, Complementary Assets and Firms'Incentives for Technology Licensing," *Management Science*, 52, 293–308.

ARORA, A., A. FOSFURI, AND A. GAMBARDELLA (2001): *Market for Technology: The Economics of Innovation and Corporate Strategy*. The MIT Press.

ARORA, A., AND A. GAMBARDELLA (1994): "The changing technology of technological change: general and abstract knowledge and the division of innovative labour," *Research Policy*, 23 (5), 523–532.

ARROW, K. (1962): "Economic Welfare and the Allocation of Resources for Invention," in *Rate and Direction of Inventive Activity*. NBER and Princeton Univ. Press, Princeton, NJ.

——— (1983): "Innovation in Large and Small Firms," in *Entrepreneurship*, ed. by J. Ronen. Lexington Books, Lexington, MA.

ASTEBRO, T. (2002): "Noncapital Investment Costs and the Adoption of CAD and CNC in U.S. metalworking industries," *RAND Journal of Economics*, 33 (4), 672–688.

BERMAN, B. (2002): "From Tech Transfer to Joint Ventures, Part 1," Cafezine.

DEGNAN, S. A. (1998): "The Licensing Payoff from U.S. R and D," *Journal of the Licensing Executives International Society*, 33 (4).

DENG, Y. (2007): "Private value of European patents," *European Economic Review*, 51 (7), 1785–1812.

DYKEMAN, D., AND D. KOPKO (2004): "Recording Patent License Agreements in the USPTO," *Intellectual Property Today*, August, 18–19.

GANS, J., D. HSU, AND S. STERN (2002): "When Does Start-Up Innovation Spur the Gale of Creative Destruction?," *RAND Journal of Economics*, 33, 571–586.

GANS, J., AND S. STERN (2003): "The Product Market and the Market for "Ideas": Commercialization Strategies for Technology Entrepreneurs," *Research Policy*, 32, 333–350.

GOFFE, W. L., G. D. FERRIER, AND J. ROGERS (1994): "Global Optimization of Statistical Functions with Simulated Annealing," *Journal of Econometrics*, 60, 65–100.

GRILICHES, Z. (1992): "The Search of R&D Spillovers," *Scandinavian Journal of Economics*, 94, 29–47.

GRINDLEY, P., AND D. J. TEECE (1997): "Managing Intellectual Capital: Licensing and Cross-licensing in Semiconductors and Electronics," *California Management Review*, 39.

HALL, B. H., A. B. JAFFE, AND M. TRAJTENBERG (2001): "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," NBER Working Paper 8498.

HALL, B. H., AND R. H. ZIEDONIS (2001): "The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995," *RAND Journal of Economics*, 32 (1), 101–128.

HARHOFF, D., F. NARIN, F. M. SCHERER, AND K. VOPEL (1999): "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, 81, 511–515.

JAFFE, A. B., AND J. LERNER (2004): *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It*. Princeton University Press, Princeton, N.J.

LAMOREAUX, N., AND K. SOKOLOFF (1997): "Inventors, Firms, and the Market for Technology: U.S. Manufacturing in the Late Nineteenth and Early Twentieth Centuries," NBER Working Paper H0098.

——— (1999): "Inventive Activity and the Market for Technology in the United States, 1840-1920," NBER Working Paper 7107.

LANJOUW, J. O. (1998): "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value," *Review of Economic Studies*, 65, 671–710.

MCFADDEN, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995–1026.

MERRILL, S. A., R. C. LEVIN, AND M. B. MYERS (2004): *A Patent System for the 21st Century*. The National Academies Press, Washington D.C.

PAKES, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54 (4), 755–784.

PAKES, A., AND D. POLLARD (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1058.

PAKES, A., AND M. SCHANKERMAN (1984): "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources," in *R&D, Patents and Productivity*, ed. by Z. Griliches, pp. 98–112. University of Chicago Press, Chicago.

PUTNAM, J. (1996): "The Value of International Patent Rights," Ph.D. thesis, Yale University.

ROBBINS, C. A. (2009): "Measuring Payments for the Supply and Use of Intellectual Property," in *International Flows of Invisibles: Trade in Services and Intangibles in the Era of Globalization*, ed. by M. Slaughter, and M. Reinsdorf. University of Chicago Press and the NBER.

ROSENBERG, N. (1996): "Uncertainty and Technological Change," in *The Mosaic of Economic Growth*, ed. by R. Landau, T. Taylor, and G. Wright. Stanford University Press.

SCHANKERMAN, M. (1998): "How Valuable is Patent Protection? Estimates by Technology Field," *RAND Journal of Economics*, 29 (1), 77–107.

SCHANKERMAN, M., AND A. PAKES (1986): "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period," *Economic Journal*, 96, 1052–1076.

SERRANO, C. J. (2006): "The Market for Intellectual Property: Evidence from the Transfer of Patents," Ph.D. thesis, University of Minnesota.

———— (2010): "The Dynamics of the Transfer and Renewal of Patents," *RAND Journal of Economics*, 41, 686–708.

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

———— (1986): "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy," *Research Policy*, 15, 285–305.

Appendix: Proofs, Tables and Figures

Lemma 1. The value function $V_a(x, y)$ is continuous and increasing in the current return of the holder of the patent, x, and the return factor of the potential buyer, y. The option value $EV_{a+1}(x', y'|x, y, d)$ is decreasing in a.

Proof. The option value $EV_{a+1}(x', y'|x, y, d)$ depends on whether or not the patent has been sold, kept, or renewed at age a. If the patent is sold $(d = 1)$, $EV_{a+1}(x', y'|x, y, d) = EV_{a+1}(x', y'|y, S)$; if the patent is kept by the current owner, $EV_{a+1}(x', y'|x, y, d) = EV_{a+1}(x', y'|x, K)$; the option value $EV_{a+1}(x', y'|x, y, d) = 0$ whenever the patent is allowed to expire. For convenience, let us consider the option value be $EV_{a+1}(x', y'|z, d)$ where $z \in \{y, x, 0\}$ depending on whether the patent was sold (y), kept (x), or allowed to expire (0). The proof is separated into two parts. Parts (i) shows that the option value is continuous and increasing with z; part (ii) proves that the option value is decreasing with a. Both proofs are by induction on a. Let L be the maximum legal length of patent protection.

Part (i). Since $E[V_{L+1}(x', y'|z)] = 0$ for all z, the initial condition of the inductive argument is satisfied. Now, suppose that the result holds for $a + 1$, it suffices to show that the result also holds for a. The hypothesis of the inductive argument is that $E[V_{a+2}(x', y'|z)]$ is continuous and increasing in z. Recall that $V_{a+1}(x, y) = \max\{V_{a+1}^S(x, y), V_{a+1}^K(x, y), 0\}$ and that $V^S(x, y) = y - c_a - \tau + \beta E[V_{a+2}(x', y'|y)]$, $V^K(x, y) = x - c_a + \beta E[V_{a+2}(x', y'|x)]$. So, $V_{a+1}(x, y)$ is also continuous and weakly increasing in $z \in \{y, x, 0\}$. Finally, we want to show that if $V_{a+1}(x, y)$ is also continuous and increasing in $z \in \{y, x, 0\}$, then $E_{g^e}[V_{a+1}(x', y'|z)]$ is also continuos and increasing in z. Since the joint distribution of the internal and external returns $F(.)$ is independent of z, we can show that if $z_1 \geq z_2$, then $E[V_{a+1}(x', y'|z_1)] \geq E[V_{a+1}(x', y'|z_2)]$.

$$
\begin{aligned}
E[V_{a+1}(x', y'|z_1)] &= \int_{u^e} \int_{u^i} V_{a+1}(u^i z, u^i z u_a^e|z_1) F_{g^i}(\mathrm{d}u^i) F_{g^i}(\mathrm{d}u^e) \\
&\geq \int_{u^e} V_{a+1}(u^i z, u^i z u_a^e|z_2) F_{g^i}(\mathrm{d}u^i) F_{g^i}(\mathrm{d}u^e) \\
&= E[V_{a+1}(x', y'|z_2)]
\end{aligned}
$$

To prove continuity, take any $z \in R^+$. $E_{g^e}[V_{a+1}(x', y'|z)]$ will be continuous at z if for every sequence $\{z_n\}$ such that $\lim(z_n) = z$, we can show that $\lim_{z_n \to z} (E[V_{a+1}(x', y'|z_n)]) = E[V_{a+1}(x', y'|z)]$. Since $F_{g^e}(.)$ and $F_{g^i}(.)$ are independent of z.

$$
\begin{aligned}
\lim_{z_n \to z} E[V_{a+1}(x', y'|z_n)] &= \lim_{z_n \to z} \int_{u^e} \int_{u^i} V_{a+1}(u^i z_n, u^i z_n u_a^e|z_n) F_{g^i}(\mathrm{d}u^i) F_{g^i}(\mathrm{d}u^e) \\
&= \int_{u^e} \int_{u^i} \lim_{z_n \to z} (V_{a+1}(u^i z_n, u^i z_n u_a^e|z_n)) F_{g^i}(\mathrm{d}u^i) F_{g^i}(\mathrm{d}u^e) \\
&= \int_{u^e} \int_{u^i} V_{a+1}(u^i z, u^i z u_a^e|z)) F_{g^i}(\mathrm{d}u^i) F_{g^i}(\mathrm{d}u^e)
\end{aligned}
$$

where second step follows, in particular, because $F(.)$ is independent of z, and the last step follows because $V_{a+1}(u^i z, u^i z u_a^e|z)$ is continuos in z (i.e., $\lim_{z_n \to z} (V_{a+1}(u^i z_n, u^i z_n u_a^e|z_n)) = V_{a+1}(u^i z, u^i z u_a^e|z)))$.

Part (ii). We want to show that $E[V_{a+1}(x', y'|z)]$ is weakly decreasing in a. The initial condition of the inductive argument requires to show that $E[V_{L+1}(x', y'|z)] \leq E[V_L(x', y'|z)]$. Since patents are active for L periods, then $E[V_{L+1}(x', y'|z)] = 0$. Moreover, by definition of $V_L(x', y'|z) \geq 0$, so it must be the case that $E[V_L(x', y'|z)]$. The induction hypothesis is that $E[V_{a+2}(x', y'|z)] \leq E[V_{a+1}(x', y'|z)]$. It suffices

36

to show that $E[V_{a+1}(x', y'|z)] \leq E[V_a(x', y'|z)]$. Recall that $V_{a+1}(x', y'|z) = \max\{y' - c_{a+1} - \tau + \beta E[V_{a+2}(x'', y''|y')], x' - c_{a+1} + \beta E[V_{a+2}(x'', y''|x')], 0\}$
$\leq \max\{y' - c_a - \tau + \beta E[V_{a+1}(x'', y''|y)], x' - c_a + \beta E[V_{a+1}(x'', y''|x')]\} = V_a(x', y'|z)$, where the inequality holds because $c_a \leq c_{a+1}$ and because of the induction hypothesis. Finally, we must show that

$$
\begin{aligned}
E[V_{a+1}(x', y'|z)] &= \int_{u^e} \int_{u^i} V_{a+1}(u^i z, u^i z u_a^e|z) F_{g^i}(du^i) F_{g^e}(du^e) \\
&\leq \int_{u^e} \int_{u^i} V_a(u^i z, u^i z u_a^e|z) F_{g^i}(du^i) F_{g^e}(du^e) \\
&= E[V_a(x', y'|z)]
\end{aligned}
$$

where the inequality follows because $V_{a+1}(x', y'|z) \leq V_a(x', y'|z)$, $F_{g^e}(.)$ is independent of a, and F_{g^i} is such that given z the $\Pr(x' > x|z, a)$ decreases with a. ∎

Lemma 2 For a fixed τ, g^e, g^i and a , we can show that: (a) If $V_a^S(x, y) - V_a^K(x, y) \geq 0$, then $V_a^S(x, y) - V_a^K(x, y)$ is increasing with x, (b) the $V_a^K(x, y) - V_a^E(x, y)$ is increasing with x.

Proof. This Lemma will be useful in the proofs of Proposition 1 and 2. For convenience, let me define $\widetilde{V}_a(x)$ as $V_a^K(x, y)$, that is

$$
\widetilde{V}_a(x) = x - c_a + \beta E[V_{a+1}(g_a^i x, g_a^i x g_a^e)]
$$

So, since $V_a^S(x, y) = V_a^K(x, y) - \tau$, then $V_a^S(x, y) = \widetilde{V}_a(y) - \tau$.

Let us start proving part (a) of the Lemma. That is the case when the owner of the patent is indifferent between selling the patent or keep it (i.e., any per period revenue such $x > \widehat{x}_a$).

The proof is by induction on a.

1) Suppose $a = L$ (i.e., the last period).

$$
V_L^S(x, y) - V_L^K(x, y) = xg^e - \tau - x = x(g^e - 1) - \tau
$$

Recall that $V_L^S(x, y) - V_L^K(x, y) \geq 0$, so it must the case that $g^e > 0$. Then, $V_L^S(x, y) - V_L^K(x, y)$ is decreasing with x.

2) Suppose that the result holds for $a + 1$. So, the induction hypothesis is that for a fixed x, the $[V_{a+1}^S(x, y) - V_{a+1}^K(x, y)]$ is increasing with x. Then, it suffices to show that for a fixed x, the $[V_a^S(x, y) - V_a^K(x, y)]$ is increasing with x. Without loss of generality we can consider that $y > x$.

We can rewrite $[V_a^S(x, y) - V_a^K(x, y)]$ as

$$
\begin{aligned}
V_a^S(x, y) - V_a^K(x, y) &= y - c_a - \tau + \beta E[V_{a+1}(g_a^i y, g_a^i y g_a^e)] \\
&\quad -x + c_a + \tau - \beta E[V_{a+1}(g_a^i x, g_a^i x g_a^e)] \\
&= (y - x - \tau) + \beta E[V_{a+1}(g_a^i y, g_a^i y g_a^e) - V_{a+1}(g_a^i x, g_a^i x g_a^e)]
\end{aligned}
$$

Since $y > x$ and $y = xg_{a-1}^e$ then It is obvious to show that $(y - x - \tau)$ is increasing with x. Let us look at the second term $E_{g^e}[V_{a+1}(g_a^i y, g_a^i y g_a^e) - V_{a+1}(g_a^i x, g_a^i x g_a^e)]$. Since F_{g^e} and F_{g^i} are independent of x, it just remain to be shown that $[V_{a+1}(g_a^i y, g_a^i y g_a^e) - V_{a+1}(g_a^i x, g_a^i x g_a^e)]$ is increasing with x.

We can rewrite the equivalent of the induction hypothesis as

$$
[V_{a+1}^S(x, y) - V_{a+1}^K(x, y)] = \widetilde{V}_{a+1}(y) - \tau - \widetilde{V}_{a+1}(x)
$$

which is weakly increasing with x.

There are three cases to study.

Case (1): fix x and consider a realization of g_a^i and g_a^e such $V_{a+1}(g_a^i y, g_a^i y g_a^e) = V_{a+1}^S(g_a^i y, g_a^i y g_a^e)$, and $V_{a+1}(g_a^i x, g_a^i x g_a^e) = V_{a+1}^K(g_a^i x, g_a^i x g_a^e)$. That is in K region with $(g_a^i x, g_a^i x g_a^e)$ and in S region with $(g_a^i y, g_a^i y g_a^e)$. Let $\lambda = g_{a-1}^e g_a^e$. Then,

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

where the last expression is an increasing transformation of the equivalent of the induction hypothesis (i.e., $\widetilde{V}_{a+1}(y) - \tau - \widetilde{V}_{a+1}(x)$). Therefore, $V_{a+1}(g_a^i y, g_a^i y g_a^e) - V_{a+1}(g_a^i x, g_a^i x g_a^e)$ is increasing with x.

Case (2): fix x and consider a realization of g_a^i and g_a^e such that, in K region with $(g_a^i y, g_a^i y g_a^e)$, in K region with $(g_a^i x, g_a^i x g_a^e)$. Let $\lambda = g_{a-1}^e$. Then,

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

where the last expression is an increasing transformation of the equivalent of the induction hypothesis (i.e., $\widetilde{V}_{a+1}(y) - \tau - \widetilde{V}_{a+1}(x)$). Therefore, $V_{a+1}(g_a^i y, g_a^i y g_a^e) - V_{a+1}(g_a^i x, g_a^i x g_a^e)$ is weakly increasing with x.

Case (3): fix x and consider a realization of g_a^e and g_a^i such that, in S region with $(g_a^i y, g_a^i y g_a^e)$ and in S region with $(g_a^i x, g_a^i x g_a^e)$. Let $\lambda = g_{a-1}^e$. Then,

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

where the last expression is an increasing transformation of the equivalent of the induction hypothesis (i.e., $\widetilde{V}_{a+1}(y) - \tau - \widetilde{V}_{a+1}(x)$). Therefore, $V_{a+1}(g_a^i y, g_a^i y g_a^e) - V_{a+1}(g_a^i x, g_a^i x g_a^e)$ is increasing with x.

Let us now prove part (b) of the Lemma. We want to show that the $V_a^K(x, y) - V_a^E(x, y)$ is increasing with x. We know that $V_a^E(x, y) = 0$. So, $V_a^K(x, y) - V_a^E(x, y) = V_a^K(x, y)$. For a fixed g^e and g_a^i, Lemma 1 shows that $V_a^K(x, y)$ is increasing with x. Therefore, $V_a^K(x, y) - V_a^E(x, y)$ is also increasing with x. ∎

Proposition 1. For each age, there exist cutoff rules $\widehat{x}_a(\theta)$ and $\widehat{g}_a^e(x, \theta) > 1$ such that patents with an improvement factor above $\widehat{g}_a^e(x, \theta)$ will be sold. Among the patent holders that met a potential buyer with improvement factor below $\widehat{g}_a^e(x, \theta)$, the patents with per period returns x above $\widehat{x}_a(\theta)$ will be renewed, and those with returns x below $\widehat{x}_a(\theta)$ will be retained by their current owners.

Proof. We first show that there exist a cutoff rule $\widehat{g}_a^e(x, \tau)$ defined as the improvement factor that makes a potential seller indifferent between whether or not to sell the patent. Two cases must be analyzed: (a) The seller

is indifferent between selling and letting the patent expire (i.e., $V_a^S(x, x\widehat{g}_a^e(x, \tau)) = 0$); and (b) The seller is indifferent between selling and keeping the patent (i.e., $V_a^S(x, x\widehat{g}_a^e(x, \tau)) = V_a^K(x, x\widehat{g}_a^e(x, \tau))$).

Let us start with case (a). Lemma 1 implies that, for a fixed a, $\widetilde{V}_a(x)$ is an increasing function of x. We can show that for any $\tau > 0$, if $g^e = 0$, $\widetilde{V}_a(xg^e) - \tau < 0$. We can also show that for any $\tau > 0$ there exist a sufficiently high g^e, for example $\overline{g^e}$, that $\widetilde{V}_a(x\overline{g^e}) - \tau > 0$. Therefore, by Bolzano's theorem, there exist an improvement factor $\widehat{g}_a^e(x, \tau)$ such that

$$\begin{aligned}
\widetilde{V}_a(x\widehat{g}_a^e(x, \tau)) - \tau &= 0 \\
V_a^S(x, x\widehat{g}_a^e(x, \tau)) &= 0
\end{aligned}$$

In case (b) the potential seller is indifferent between selling and keeping the patent. That is $V_a^S(x, y) = V_a^K(x, y)$, which implies that

$$\widetilde{V}_a(x\widehat{g}_a^e(x, \tau)) - \tau = \widetilde{V}_a(x)$$

Now, let us consider a $\tau > 0$ and a $g^e = 1$. We can show that $V_a^S(x, y) - V_a^K(x, y) < 0$ because $V_a^S(x, y) = V_a^S(x, x) - \tau = V_a^K(x, y) - \tau$. On the other hand, for fixed τ and fixed x, Lemma 1 shows that $V_a^S(x, y)$ is increasing in y. So, for sufficiently high g^e we can show that $V_a^S(x, y) - V_a^K(x, y) > 0$. Finally, by Lemma 2 we know that for a fixed τ, the difference $(V_a^S(x, y) - V_a^K(x, y))$ is increasing in x. By Bolzano's theorem, there exist a $\widehat{g}_a^e(x, \tau)$ such that $V_a^S(x, x\widehat{g}_a^e(x, \tau)) = V_a^K(x, x\widehat{g}_a^e(x, \tau))$, which is that

$$\widetilde{V}_a(x\widehat{g}_a^e(x, \tau)) - \tau = \widetilde{V}_a(x)$$

The second part of the proof shows that there exist a per period return $\widehat{x}_a(\theta)$ such that patents with per period return $x < \widehat{x}_a(\theta)$ will be allowed to expire while the rest of the patents will be renewed. The cutoff rule $\widehat{x}_a(\theta)$ is defined as the patent revenue that makes a patent owner indifferent between keeping the patent and letting it expire (i.e., $V_a^K(\widehat{x}_a, y) = 0$). Recall that $V_a^K(x, y) = x + \beta E[V_{a+1}(x', y')|x, y, K] - c_a$ and the function $V_a^K(x, y)$ is increasing with x by Lemma 1. Moreover, if $x = 0$, $V_a^K(x, y) < 0$; and if x is positive and sufficiently large $V_a^K(x, y) > 0$. Therefore, Bolzano's theorem implies that there is a $\widehat{x}_a$ such that $V_a^K(\widehat{x}_a, y) = 0$). ∎

Proposition 2. The following properties characterize the functions $\widehat{x}_a(\theta)$ and $\widehat{g}_a^e(x, \theta)$:

    (a) The function $\widehat{x}_a(\theta)$ is increasing with patent age a at the renewal dates.

    (b) If the costs of technology transfer τ are positive, then

        (b.1) For a fixed patent age a, the function $\widehat{g}_a^e(x, \theta)$ is monotonically decreasing with x

        (b.2) For a fixed patent return x, the function $\widehat{g}_a^e(x, \theta)$ is monotonically increasing with patent age a

Proof. Part (a). We want to show that $\widehat{x}_a(\theta)$ is increasing with patent age a at the renewal dates. From proposition 1 we know that that there exist a renewal date a and patent revenue $\widehat{x}_a$ such that $V_a^K(\widehat{x}_a, \widehat{x}_a g^e) = 0$. Similarly we can obtain that $V_{a+1}^K(\widehat{x}_{a+1}, \widehat{x}_{a+1} g^e) = 0$ because the schedule of renewal fees is increasing with a. Note that $V_{a+1}^K(\widehat{x}_{a+1}, \widehat{x}_{a+1} g^e) = V_a^K(\widehat{x}_a, \widehat{x}_a g^e)$. Because the function $V_a^K(., .)$ is increasing with a and x (see Lemma 1). then it must be the case that $\widehat{x}_{a+1} \geq \widehat{x}_a$ for the above equality to hold.

Part (b.1) [Selection Effect]. We want to show that the cutoff rule $\widehat{g}_a^e(x, \tau)$ is decreasing with x. For convenience, let me define $V_a^K(x, y)$ as $\widetilde{V}_a(x)$, that is

$$\widetilde{V}_a(x) = x - c_a + \beta E[V_{a+1}(g_a^i x, g_a^i x g_a^e)]$$

So, since $V_a^S(x, y) = V_a^K(x, y) - \tau$, then $V_a^S(x, y) = \widetilde{V}_a(y) - \tau$.

Let us first consider when the owner of the patent is indifferent between selling the patent or let it allow to expire. That is any per period revenue such $x \leq \widehat{x}_a$. Proposition 1 implies that there is a cutoff rule $\widehat{g}_a^e(x, \tau)$ and that is defined as follows.

$$
\begin{aligned}
V_a^S(x, x\widehat{g}_a^e(x, \tau)) &= 0 \\
\widetilde{V}_a(x\widehat{g}_a^e(x, \tau)) - \tau &= 0
\end{aligned}
$$

From Lemma 1 we know that for a fixed g^e the function $V_a^S(x, xg^e) = \widetilde{V}_a(xg^e) - \tau$ is increasing with x. So, for the above equality to hold, it must be the case that if x increases then $\widehat{g}_a^e(x, \tau)$ decreases.

Let us now consider the case when the owner of the patent is indifferent between selling the patent or keep it. That is any per period revenue such $x > \widehat{x}_a$.

The proof is by induction on a.

1) Suppose $a = L$ (i.e., the maximum legal length of patent protection). A seller is indifferent between selling and keeping a patent if

$$
\begin{aligned}
V_L^S(x, y) &= V_L^K(x, y) \\
x\widehat{g}_a^e(x, \tau) - \tau &= x
\end{aligned}
$$

And we can show that $\widehat{g}_a^e(x, \tau)$ is decreasing with x

$$
\widehat{g}_a^e(x, \tau) = 1 + \frac{\tau}{x}
$$

2) Suppose now that $\widehat{g}_a^e(x, \tau)$ is decreasing with x for $a + 1$. We want to show that the result holds for a. Notice that the induction hypothesis (i.e., $\widehat{g}_a^e(x, \tau)$ is decreasing with x for $a+1$) is equivalent to showing that for a fixed x, the $[V_{a+1}^S(x, y) - V_{a+1}^K(x, y)]$ is weakly increasing in x. Therefore, in order to prove that $\widehat{g}_a^e(x, \tau)$ in decreasing in x for age a, it suffices to show that for a fixed x, the difference $[V_a^S(x, y) - V_a^K(x, y)]$ is increasing in x.

Note that $\widehat{g}_a^e(x, \tau)$ for any a is only defined for x such that $\widehat{g}_a^e(x, \tau) \geq 1 \ \forall a \in [1, L]$. So, without loss of generality we can consider $y > x$. Finally, by Lemma 2, see especially part (2) of the proof, we show that for a fixed τ, g^e, and g^i the $[V_a^S(x, y) - V_a^K(x, y)]$ is increasing with x. As argued above, this results is equivalent to showing that the function $\widehat{g}_a^e(x, \tau)$ is decreasing with x.

Part (b.2) [Horizon Effect]. We want to show that for any $x > 0$ and $\tau > 0$, the cutoff rule $\widehat{g}_a^e(x, \tau)$ is increasing with a. For convenience, let me define $\widetilde{V}_a(x)$ as $V_a^K(x, y)$ as

$$
\widetilde{V}_a(x) = x - c_a + \beta E_{g^e}[V_{a+1}(\delta x, \delta x g_a^e)]
$$

So, since $V_a^S(x, y) = V_a^K(x, y) - \tau$, then $V_a^S(x, y) = \widetilde{V}_a(y) - \tau$.

Showing that for a fixed x and τ the cutoff rule $\widehat{g}_a^e(x, \tau)$ is weakly increasing with a is equivalent to proving that for a fixed x and g^e the difference $[V_a^S(x, y) - V_a^K(x, y)]$ is decreasing with a. That is,

$$
[V_a^S(x, y) - V_a^K(x, y)] > [V_{a+1}^S(x, y) - V_{a+1}^K(x, y)]
$$

And rearranging the induction hypothesis we obtain,

$$\widetilde{V}_a(y) - \tau - \widetilde{V}_a(x) > \widetilde{V}_{a+1}(y) - \tau - \widetilde{V}_{a+1}(x)$$

$$\widetilde{V}_a(y) - \widetilde{V}_a(x) > \widetilde{V}_{a+1}(y) - \widetilde{V}_{a+1}(x)$$

$$y - c_a + \beta E[V_{a+1}(g_a^i y, g_a^i y g_a^e)] - x + c_a - \beta E[V_{a+1}(g_a^i x, g_a^i x g_a^e)] > y - c_{a+1} + \beta E[V_{a+2}(g_{a+1}^i y, g_{a+1}^i y g_{a+1}^e)]$$
$$- x + c_{a+1} - \beta E[V_{a+2}(g_{a+1}^i x, g_{a+1}^i x g_{a+1}^e)]$$

$$y - x + \beta E[V_{a+1}(g_a^i y, g_a^i y g_a^e)] - \beta E[V_{a+1}(g_a^i x, g_a^i x g_a^e)] > y - x + \beta E[V_{a+2}(g_{a+1}^i y, g_{a+1}^i y g_{a+1}^e)]$$
$$- \beta E[V_{a+2}(g_{a+1}^i x, g_{a+1}^i x g_{a+1}^e)]$$

$$E[V_{a+1}(g_a^i y, g_a^i y g_a^e)] - E[V_{a+1}(g_a^i x, g_a^i x g_a^e) > E[V_{a+2}(g_{a+1}^i y, g_{a+1}^i y g_{a+1}^e)]$$
$$- E[V_{a+2}(g_{a+1}^i x, g_{a+1}^i x g_{a+1}^e)]$$

Without loss of generality, consider per period revenues x and y, $y > x$, such that $[V_a^S(x,y) - V_a^K(x,y)] \geq 0$.

The proof is by induction.

1) Suppose $a + 1 = L$, so $a = L - 1$. We want to show that

$$[V_L^S(x,y) - V_L^K(x,y)] > [V_{L+1}^S(x,y) - V_{L+1}^K(x,y)]$$

Let us start defining

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

$$V_{L-1}^S(x,y) - V_{L-1}^K(x,y) = y + \beta E[V_L(g_{L-1}^i y, g_{L-1}^i y g_{L-1}^e)] - \tau - x + \beta E[V_L(g_{L-1}^i x, g_{L-1}^i x g_{L-1}^e)]$$
$$= y - \tau - x + \beta E[V_L(g_{L-1}^i y, g_{L-1}^i y g_{L-1}^e) - V_L(g_{L-1}^i x, g_{L-1}^i x g_{L-1}^e)]$$

In Lemma 1 we showed that for a fixed a, $V_a(x,y)$ was weakly increasing in x. So, it must be the case that $V_L(g_{L-1}^i y, g_{L-1}^i y g_{L-1}^e) \geq V_L(g_{L-1}^i x, g_{L-1}^i x g_{L-1}^e)$. The expectation of a random variable that is larger or equal than zero is also larger or equal than zero. So, $E[V_L(g_{L-1}^i y, g_{L-1}^i y g_{L-1}^e) - V_L(g_{L-1}^i x, g_{L-1}^i x g_{L-1}^e)] \geq 0$. Then

$$V_{L-1}^S(x,y) - V_{L-1}^K(x,y) = y - \tau - x + \beta E[V_L(g_{L-1}^i y, g_{L-1}^i y g_{L-1}^e) - V_L(g_{L-1}^i x, g_{L-1}^i x g_{L-1}^e)]$$
$$\geq y - \tau - x$$
$$= V_L^S(x,y) - V_L^K(x,y)$$

Therefore, $[V_L^S(x,y) - V_L^K(x,y)] > [V_{L+1}^S(x,y) - V_{L+1}^K(x,y)]$.

2) Now, suppose that the relationship holds for $a + 1$. So, the induction hypothesis is

$$[V_{a+1}^S(x,y) - V_{a+1}^K(x,y)] > [V_{a+2}^S(x,y) - V_{a+2}^K(x,y)]$$

$$\widetilde{V}_{a+1}(y) - \tau - \widetilde{V}_{a+1}(x) > \widetilde{V}_{a+2}(y) - \tau - \widetilde{V}_{a+2}(x)$$

$$\widetilde{V}_{a+1}(y) - \widetilde{V}_{a+1}(x) > \widetilde{V}_{a+2}(y) - \widetilde{V}_{a+2}(x)$$

We will show that it holds for a, that is

$$[V_a^S(x,y) - V_a^K(x,y)] > [V_{a+1}^S(x,y) - V_{a+1}^K(x,y)]$$

We know that

$$
\begin{aligned}
V_a^S(x,y) - V_a^K(x,y) &= \widetilde{V}_a(y) - \tau - \widetilde{V}_a(x) \\
&= y - \tau + \beta E[V_{a+1}(g_a^i y, g_a^i y g_a^e)] - x - \beta E[V_{a+1}(g_a^i x, g_a^i x g_a^e)] \\
&= y - \tau - x + \beta(E[V_{a+1}(g_a^i y, g_a^i y g_a^e)] - E[V_{a+1}(g_a^i x, g_a^i x g_a^e)]) \\
&> y - \tau - x + \beta(E[V_{a+2}(g_{a+1}^i y, \delta y g_{a+1}^e)] - E[V_{a+2}(g_{a+1}^i x, \delta x g_{a+1}^e)]) \\
&= y - \tau - c_a + \beta E[V_{a+2}(g_{a+1}^i y, g_{a+1}^i y g_{a+1}^e)] - x + c_a - \beta E[V_{a+2}(g_{a+1}^i x, g_{a+1}^i x g_{a+1}^e)] \\
&= V_{a+1}^S(x,y) - V_{a+1}^K(x,y)
\end{aligned}
$$

∎

Table A-1: Distribution of the Patent Returns at Age One, Four, and Seven

Patent Returns	Age 1 (1)	Age 4 (2)	Age 7 (3)
50	0.6	9.7	22.1
250	5.2	15.7	25.6
500	11.1	22.6	33.2
1,000	20.6	32.9	44.0
5,000	54.7	64.4	72.3
10,000	69.8	77.1	82.8
25,000	85.2	89.3	92.5
100,000	96.8	97.8	98.6
500,000	99.7	99.9	99.9
Mean	17,999	12,716	9,019

Table A-2: Distribution of the Value of Patents at Age 1, 4, and 7

	Patent value at age 1		Patent value at age 4		Patent value at age 7	
Percentile	Value (US$ 2003) (1)	Cum % of Total	Value (US$ 2003) (2)	Cum % of Total	Value (US$ 2003) (3)	Cum % of Total
5	787	0.0	0	0.0	0	0.0
10	1,472	0.1	134	0.0	0	0.0
25	4,584	0.7	1,426	0.2	431	0.0
50	16,184	3.7	8,664	2.3	4,180	1.2
75	54,950	13.9	35,616	11.2	22,231	8.8
80	74,198	18.1	48,881	15.2	31,379	12.5
90	160,829	32.3	110,991	29.1	75,121	26.1
98.5	763,483	67.5	550,287	65.4	390,948	63.3
99.8	2,432,580	88.0	1,777,078	86.9	1,303,508	86.2
Mean	76,598		52,669		35,835	

Table A-3: Distribution of the Value of Traded and Untraded Patents

Percentile	Untraded patents (1)		Traded patents (2)	
	Value (US$ 2003)	Cum % of Total	Value (US$ 2003)	Cum % of Total

A. Distribution of the value of patents

Percentile	Value (US$ 2003)	Cum % of Total	Value (US$ 2003)	Cum % of Total
5	654	0.0	4,877	0.0
10	1,184	0.1	9,561	0.3
25	3,353	0.8	23,948	1.8
50	10,605	4.0	58,320	7.7
75	33,955	13.7	149,614	22.1
80	44,978	17.6	188,927	27.2
90	98,251	30.8	355,898	42.7
98.5	502,777	64.8	1,285,141	74.6
99.8	1,687,702	87.0	3,754,001	90.6
Mean value	50,162		164,670	

B. Summary statistics

Percentage (of all patents)	0.77	0.23
Value of patents as a percentage of the total value	0.51	0.49

Table A-4: Probability that a Patent is Traded over the Life Cycle

Percentile Initial return at age one	Probability of Trade $[\tau = 5,495]$ (1)	Probability of Trade $[\tau = 2,747]$ (2)
<5	0.020	0.020
5-10	0.034	0.034
10-25	0.048	0.064
25-50	0.122	0.215
50-75	0.288	0.393
75-80	0.397	0.476
80-90	0.455	0.514
90-98.5	0.528	0.559
98.5-99.8	0.535	0.562
All sample	0.231	0.296

Table A-5: Distribution of the Gains from Trade of Traded Patents

	Gains from trade	
		Cum %
Percentile	Value	of Total
5	17	0.0
10	134	0.1
25	972	0.7
50	3,416	3.7
75	10,829	13.9
80	14,357	18.1
90	30,970	32.3
98.5	148,129	67.5
99.8	540,432	88.0
Mean value	15,008	

Table A-6: Value of Traded Patents and Gains from Trade in the Market for Patents

Value of Traded Patents With and Without Option to Sell the Patent and Gains from Trade $[\tau = \$5,495]$

Percentile Initial returns at age one of traded patents	Market for patents active Average Patent value (US\$ 2003) (1)	Market for patents non-active Average Patent value (US\$ 2003) (2)	Gains from Trade (%) (3)
<5	466	466	0.000
5-10	1,190	1,190	0.000
10-25	3,066	3,038	0.010
25-50	10,839	10,283	0.054
50-75	34,380	31,577	0.089
75-80	66,073	59,933	0.102
80-90	112,510	102,167	0.101
90-98.5	328,650	298,253	0.102
98.5-99.8	1,250,645	1,119,970	0.117
Mean (all)	164,670	149,662	0.100

Table A-7: Percentage Changed in the Gains from Trade of Traded Patents and Cummulative Gains from Trade After Decreasing the Costs of Technology Transfer by Fifty Percent

Percentile Initial return at age one	Change in the Gains from Trade (1)	Additional Average Gains from Trade $[\tau = 2,747]$ (Cum %) (2)	Average Gains from Trade $[\tau = 5,495]$ (cum %) (3)
<25	6.662	0.004	0.000
25-50	1.397	0.068	0.005
50-75	0.433	0.306	0.063
75-80	0.239	0.385	0.098
80-90	0.167	0.598	0.234
90-98.5	0.062	0.829	0.628
98.5-99.8	0.043	0.935	0.890
All patents	0.106		

Figure 2: Probability that an active patent is being allowed to expire for previously traded and untraded patents
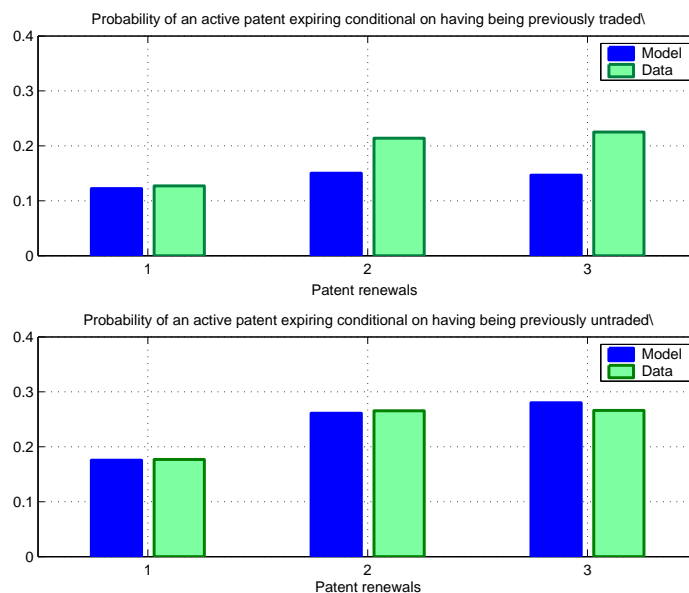

Figure 3: Probability that an active patent is traded for previously traded and untraded patents