ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

KNOWLEDGE FLOW AND SEQUENTIAL INNOVATION: IMPLICATIONS FOR TECHNOLOGY DIFFUSION, R&D AND MARKET VALUE

Sharon Belenzon

Number 259

March 2006

Manor Road Building, Oxford OX1 3UQ

Knowledge Flow and Sequential Innovation: Implications for Technology Diffusion, R&D and Market Value

Sharon Belenzon*

February 2006

Abstract

It is shown that spillovers can enhance private returns to innovation if they feed back into the dynamic research of the original inventor (Internalized spillovers), but will always reduce private returns, if the original inventor does not benefit from the advancements other inventors build into the "spilled" knowledge (Externalized spillovers). I empirically identify unique patterns of knowledge flows (based on patent citations), which provide information about whether "spilled" knowledge is reabsorbed by its inventor. A simple model of sequential innovation with dynamic spillovers is developed, which predicts that market value and R&D expenditures should rise with Internalized spillovers and fall with Externalized spillovers. These predications are confirmed using panel data on U.S. firms between 1981 and 2001. To the extent that firms internalize some of the spillovers they create, the classical underinvestment problem in R&D will be mitigated and the central role of spillovers in promoting economic growth will be enhanced.

Keywords: market value, patents, R&D and spillovers
JEL Classification: O31, O32 and O33

Acknowledgement: This was the main chapter of my PhD thesis. I deeply appreciate the tremendous support of my PhD advisors Mark Schankerman and John Van Reenen. I thank Manuel Trajtenberg, Nick Bloom, Steve Bond, Bronwyn Hall, Iain Cockburn, John Haltiwanger, Sam Kortum, Paul Klemperer, Steve Redding, John Sutton and numerous seminar participants for helpful comments.
Correspondence: sharon.belenzon@nuffield.ox.ac.uk

*Dept. of Economics, Oxford University, Nuffield College, Manor Road Building, Manor Road, Oxford, OX1 3UQ, UK.

1. Introduction

Knowledge spillovers are at the centre of the modern endogenous growth literature. Numerous studies have analyzed, empirically and theoretically, the contribution of knowledge spillovers for economic performance[1]. The present paper studies knowledge spillovers (hereafter, spillovers) in the context of sequential innovation: 'spilled' knowledge inspires follow-up research outside the boundaries of the inventing firm.

It has been well recognized that spillovers have countervailing implications for the incentive to innovate: on the one hand, spillovers encourage future research, but on the other hand, they discourage current research due to obsolescence. There is an essential distinction between the obsolescence of knowledge and the obsolescence of the rents an inventor captures on its knowledge. When knowledge is sequentially developed it cannot become obsolete since it inspires subsequent developments and it is embodied in a more advanced knowledge. However, the stream of rents the original inventor captures on its old knowledge could depreciate, especially if all the subsequent developments are done by other agents.

Schumpeter (1943) introduced the concept of creative destruction; the arrival of new knowledge renders old knowledge obsolete[2]. For example, the arrival of knowledge of how to produce power using the steam engine renders the knowledge of producing power using water obsolete. Yet, innovations that render knowledge obsolete cannot be sequential. For example, the knowledge of converting heat to work has been embodied in the first Newcome automobile engine and later in the vapor pressure engine, before being substantially improved by Englishman who studied the conversion rates from heat and back, which inspired a whole field of "thermodynamics" research[3]. The original knowledge embodied

[1] The theory of endogenous growth was pioneered by Romer (1986) and Grossman and Helpman (1991). The empirical literature has studied various types of spillovers. Griliches (1992) provides a comprehensive survey of the micro empirical literature.

[2] Creative destruction both limits and generates the incentive to innovate by rent seeking firms. This idea is modeled in various ways in the endogenous growth literature. The obsolescence of private returns to R&D has been empirically addressed, using patent renewals, by Schankerman and Pakes (1986), Pakes (1986) and Lanjouw (1998).

[3] Based on Mokyer(2005).

in the Newcome engine did not become obsolete; on the contrary, it was the basis of many subsequent improvements. Yet, its commercial value to Newcome should have become obsolete once the original knowledge spread to other agents and was sequentially advanced by them.

However, should we still argue that private returns to the old knowledge become obsolete, if the inventing firm reabsorbs in a future period its old knowledge including all the advancements the other agents have built into it? I empirically show that the answer should be *no*. Spillovers that feed back into the inventing firm should mitigate the negative effect spillovers have on private returns, since in this case, the inventing firm can still extract rents from its old knowledge even though other agents have further advanced it.

In their classic paper, Aghion and Howitt (1992) model the idea that spillovers raise private obsolescence as *a negative dependency of current research on future research*. Accordingly, the inventing firm has a lower incentive to innovate in case its knowledge is spilled to other agents that further advance it. Actually, spillovers can generate a "no-growth trap"; the negative effect of spillovers on private obsolescence can be so large that the incentive to create the first generation of knowledge completely disappears.

Nonetheless, in a dynamic context, this negative effect of spillovers could be substantially mitigated. The inventing firm could reabsorb its spilled knowledge in a future period, which should reduce the obsolescence of rents the inventing firm captures. Suppose there are two economies that are identical in all dimensions, but in the first economy firms are more likely to reabsorb their spilled knowledge. We would expect innovation and growth to be higher in the first economy, since spillovers would discourage current research to a lesser extent.

The major contribution of this paper is to develop an empirical methodology, based on patents and citations, that allows measuring spillovers and the extent they feed back into the inventing firm. I define this feeding back of spillovers as *technological internalization* and show it is an important channel through which private rents are appropriated. Technological internalization is identified by distinguishing between two types of spillovers:

Internalized and *Externalized*. Internalized spillovers are spillovers that feed back into the dynamic research of the inventing firm, whereas Externalized spillovers do not. Technological internalization is argued to be higher when spillovers are more Internalized and less Externalized.

In addition to the technological channel of internalization, firms can internalize rents through a contractual channel. The literature has studied the theoretical aspects of contractual internalization, mainly as a mechanism through which rents are shared between early innovators and their followers. Green and Scotchmer (1995), Scotchmer (1996) and Chang (1995) study the theoretical aspects of the effect of a second-generation invention on the rents captured on the first-generation invention[4].

However, the literature has not yet investigated the technological channel through which an inventor can reap the rents on its discovery. Potentially, this channel of internalization could be highly important for the generation of "pure" spillovers. According to the endogenous growth literature, "pure" spillovers, which occur when knowledge transfers freely across inventors, allow the economy to depart from decreasing returns in the production of knowledge and achieve sustained economic growth. Contractual internalization hinders the free access to knowledge (since the receivers of knowledge have to incur some usage costs). Hence, contractual internalization should diminish economic growth, through restricting the increasing returns in knowledge production. Yet, under technological internalization, "pure" spillovers should not diminish in any obvious way, since private rents can be captured without limiting future research. Thus, technological internalization should be a more desirable channel through which private rents are captured.

A famous example that illustrates the importance of technological internalization for private returns is the invention of the CT (Computed Tomography) scanners. Trajtenberg (1990) finds that this invention is associated with large social returns. However, private returns to this discovery were low, as the spillovers this invention created were mostly Externalized. This technology was developed by EMI (a British electronic company) and

[4]Bessen and Maskin (2002) argue that competition in research could actually encourage the incentive to innovate when innovation is cumulative and sequential.

was patented in 1973. A market for CT scanners emerged following rapid innovation aiming at exploiting the new technological opportunities. From 1975 onwards, hundreds of offspring inventions were created. In the early years, most of these inventions were developed by EMI itself, however, after less than a decade EMI failed to capture any significant portion of the market (which was mainly dominated by General Electric) and was no longer at the frontier of the technology it had originated. This implies that the private returns to the invention of the CT scanners were strongly negatively affected by the spillovers this invention created.

The essence of my empirical methodology is as follows: knowledge is identified as a patent and knowledge flow is identified as a patent citation[5]. For each patent in the sample a "family-tree" is constructed, based on the citations the patent receives. Figure 1 illustrates this methodology for a simple case of a sequence of three patents. Assume patent m cites patent l and patent n cites patent m. Hence, the "family-tree" of patent l includes both patent m and patent n, where, patent m is the 'child' of patent l and patent n is the 'grandchild' of patent l. Given this "family-tree", invention n is classified as an offspring of invention l, even though knowledge did not transfer directly from invention l to invention l. Applying this method to a high-order sequence of citations allows tracing the trajectory knowledge has followed, while spreading across inventions and firms. Based on these trajectories, it can be determined whether knowledge that leaves the inventing firm and is further advanced by other firms will have been reabsorbed by the inventing firm in a future period (e.g., if patents l and n are held by the same firm whereas patent m is owned by another firm, the spillovers created by invention l are technologically internalized by the inventing firm).

[5]Prior studies that empirically identified citations as knowledge flows are Jaffe, Henderson and Trajtenberg (1993), Caballero and Jaffe (1993) and Jaffe and Trajtenberg (1999).

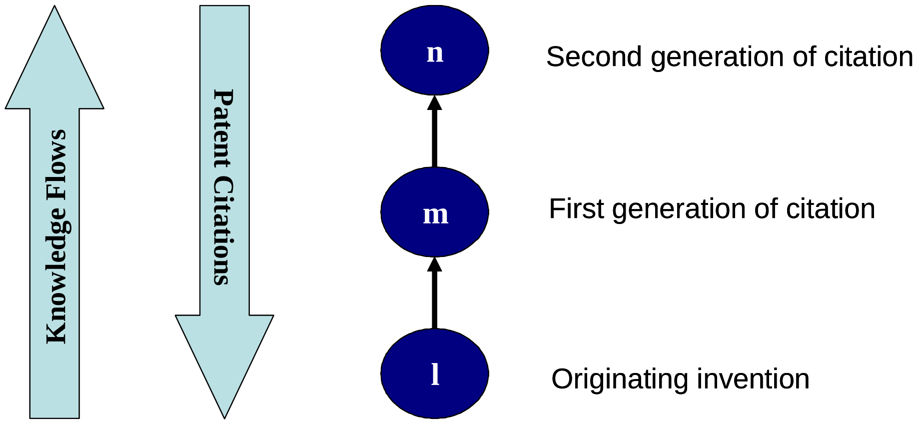


Figure 1: The "family-tree" of invention l

A model of sequential innovation with dynamic spillovers is developed: private returns to the knowledge k are defined as the stream of profits the inventing firm captures from all the subsequent developments it introduces along the line of research k originates. Spillovers are modeled as the follow-up developments of k that are invented by other agents along the line of research. In a dynamic perspective, private rents and, therefore, the R&D expenditures of the inventing firm would depend on the extent the inventing firm is able to build on the external developments of k.

Thus, the model predicts that technological internalization would raise the private returns and R&D expenditures of the inventing firm. These predictions are confirmed using panel data on the largest 500 patenting firms in the US between 1980 and 2001. There is a substantial firm-level variation in technological internalization, even within narrowly defined industries. This firm–level variation is exploited in estimating the effect of technological internalization on the market valuation of the R&D stock of the firm. The estimation results show that the effect of the R&D stock on market value intensifies when technological internalization is higher. A one standard deviation increase in the measure of Internalized spillovers raises, at the mean, the market valuation of an additional dollar spent on R&D by 30 percent, whereas a one standard deviation increase in the measure of Externalized spillovers lowers, at the mean, the market valuation of an additional dollar spent on R&D by 10 percent.

In addition to quantifying the effect of technological internalization on private returns, the findings from the market value estimation also suggest that firms themselves are aware of their technological internalization and take it into consideration when making the R&D decisions. This hypothesis is confirmed by estimating a R&D equation; firms that create more Internalized and less Externalized spillovers, on average, invest substantially more in R&D.

Finding that the R&D decision of the inventing firm is affected by the type of spillovers it creates, once more, has important implications for the endogenous growth literature. Spillovers encourage the innovation activity of the receivers of knowledge, however, the effect of spillovers on the incentive to create the spilled knowledge at the first place depends on whether they are Internalized or Externalized.

In summary, I show that firms are able to internalize dynamically some of their knowledge that spills to other firms. To the extent that such internalization occurs, the classical underinvestment problem in R&D will be mitigated, as the negative effect of spillovers on private returns weakens.

The rest of the paper is organized as following: section 2 presents the analytical framework, section 3 describes the empirical methodology, section 4 discusses the data, the econometric specifications and findings are reported in section 5 and section 6 concludes.

2. Analytical motivation

Spillovers enhance the technological opportunities created by knowledge by increasing the probability that subsequent developments of the knowledge will occur. The extent spillovers raise private returns depends on whether the inventing firm benefits from these enhanced technological opportunities. Technological internalization should be higher when spillovers are more likely to feed back into the inventing firm.

Technological internalization is represented by the parameter θ. It is assumed that firms do not behave in any strategic way to affect θ or the flow of their knowledge to other firms (although these might be interesting extensions). Also, the inventing firm cannot extract monetary payoffs from the spread of its knowledge to follow-up developers (i.e., contractual internalization is not allowed). The model does not include time. Periods are defined by the arrival of stages of developments (i.e., every period includes one stage of development) and there is no discounting.

The model distinguishes between *static* and *dynamic* returns to innovation[6]. Static returns are defined as the one period stream of profits attributed to a single invention[7]. Dynamic returns, however, also consider the stream of profits the inventing firm can capture by continuing to invent along the line of research it originates. The ability to do so depends on the extent the inventing firm can build on the follow-up developments that other firms introduce along the line of research.

Suppose firm i (the inventing firm) holds a piece of knowledge k, which has the potential of being sequentially developed an infinite number of times. The static returns to this knowledge include the stream of profits firm i receives from this invention, until these profits become obsolete, which occurs with the development of the next generation of the knowledge k. Nevertheless, dynamic returns to the knowledge k do not become obsolete once a subsequent development takes place, if firm i continues to invent along the line of research k originates.

Let v be a constant one-shot pay-off associated with winning a development stage of the knowledge k. In every development stage there are n firms competing in a patent race. Only one firm is allowed to win a development stage. If more than one firm makes a discovery, a patent cannot be granted, and both firms engage in Bertrand competition that drives profits to zero.

Every generation of development requires a constant R&D investment of x, which yields a positive probability p of making a discovery by firm i (i.e., with probability p firm

[6]Bessen and Maskin (2002) also provide a model of sequential innovation that studies the incentive to innovate in a similar dynamic framework.

[7]The huge literature on the optimal design of patents rights studies the effect of the outward flow of knowledge on static private returns (which patents aim to protect). See, for example, Klemperer (1990), Gilbert and Shapiro (1990), Scotchmer (1999) and Cornelli and Schankerman (1999).

i discovers a piece of knowledge that awards the static payoff v, if no other firm invents the given generation). Denote by q the positive probability that at least one of firm i's competitors along the line of research makes a discovery in every stage of development.

Hence, the expected static rent firm i captures from participating in a development stage, Z, which is assumed to be strictly positive, is:

$$Z = p(1-q)v - x \tag{2.1}$$

2.1. Dynamic returns and reabsorbing spilled knowledge

As a departure point, assume that once firm i fails to win in a given generation, it cannot continue developing the next generation, even if some other firm has been successful in inventing this generation. In this case, technological internalization cannot occur. Dynamic returns to the knowledge k are:

$$W_i = (p(1-q)v - x) + p(1-q)\left(p(1-q)v - x\right) + p^2(1-q)^2\left(p(1-q)v - x\right) + ... \tag{2.2}$$

The first term on the right hand side of equation (2.2) is the expected static pay-off of winning the first generation of development, the second term is the expected static pay-off of winning the second generation of development, which is positive with probability $p(1-q)$ (the probability of winning the first generation), the third term is the expected static pay-off of winning the third generation (which is positive only if firm i had won the first and second generations of developments that occur with probability $p^2(1-q)^2$) and so forth. It is straightforward to show that since an infinite number of potential developments is assumed, W_i becomes:

$$W_i = \frac{vp(1-q) - x}{1 - p(1-q)} \tag{2.3}$$

Next, suppose that if firm i does not win in a given generation, whereas the subsequent knowledge has been invented, firm i can still proceed to invent the follow-up generation.

Thus, the extent technological internalization occurs is captured by the number of 'second chances' the inventing firm gets to stay in the sequential development of its spilled knowledge, if it fails to win in a development stage, but some other firm invents. The number of "second chances" is denoted by θ (in equation (2.3), $\theta = 0$, since the firm is not allowed to have any 'second chances').

Consider the case where $\theta = 1$, i.e., if firm i fails to win more than once, it is forced out from the dynamic race (firm i receives one 'second chance'). In this scenario, the dynamic returns firm i captures on its knowledge k can be written as:

$$\begin{aligned} W_i(\theta &= 1) = (p(1-q)v - x) + p(1-q)\left(p(1-q)v - x\right) + p^2(1-q)^2\left(p(1-q)v - x\right) + ... \\ & +p(1-q)q(p(1-q)v - x) + 2p(1-q)(1-p)q\left(p(1-q)v - x\right) + ... \end{aligned} \tag{2.4}$$

Where the second row on the right hand side of equation (2.4) represents the 'second chance' firm i gets (for example, the first term in the second row is the additional expected rent the firm captures due to the fact it is allowed not to win in the first generation and still participate in the development race of the second generation). It is easy to show that equation (2.4) can be written as:

$$W_i(\theta = 1) = \frac{(1-q)\,pv - x}{(1-p)\,(1-q)}\left(1 - \left(\frac{q}{1-(1-q)\,p}\right)^2\right) \tag{2.5}$$

This model can be generalized for any θ in the following way[8]:

$$W_i(\theta) = \frac{(1-q)\,pv - x}{(1-p)\,(1-q)}\left(1 - \left(\frac{q}{1-(1-q)\,p}\right)^{\theta+1}\right) \tag{2.6}$$

This implies that the returns firm i faces increase in θ, since $\frac{q}{1-(1-q)p} < 1$[9]. This summarizes the main theoretical prediction; accordingly, private returns rise with the ability to reabsorb spilled knowledge (which increases the likelihood of technological internalization).

[8] See Belenzon (2005) for detail on the derivation of this expression.

[9] q is also the probability an invention occurs, however, firm i is not the winner. Thus, $q = q(1-p)+pq$, which is smaller than $1 - p + pq$, as $q < 1$.

Note that by substituting $\theta = 0$ into equation (2.6) we get equation (2.3), which is the dynamic returns firm i captures on its discovery without technological internalization. Moreover, when the firm has a 'complete' ability to build on the research of its rivals ($\theta = \infty$), the dynamic returns become:

$$W_i(\theta = \infty) = \frac{v\,(1-q)\,p - x}{(1-p)\,(1-q)} \tag{2.7}$$

Hence, dynamic returns are the static returns per subsequent invention (Z in equation (2.1)), discounted by the probability that the line of research will terminate[10] (which occurs with the probability $(1-p)\,(1-q)$, where no firm invents). As the probability that the line of research terminates falls, dynamic returns rise.

2.2. The incentive to innovate and θ

θ should positively affect the innovation effort of the inventing firm since it increases private returns (the subscript i is omitted in the rest of this section).

Suppose k originates two different lines of research. In the first line of research $\theta = 0$ and in the second line of research $\theta = \infty$. The next section shows how these two types of lines of research are empirically identified. Lines of research with $\theta = 0$ are associated with knowledge leaving the boundaries of the inventing firm, never to return (defined as Externalized lines of research), whereas lines of research with $\theta = \infty$ are associated with the originating knowledge being reabsorbed by the inventing firm (defined as Internalized lines of research).

The model is slightly modified to allow firm i choosing its R&D expenditures, x, which affects the probability of inventing, $p(x)$, with $p'(x) > 0$, $p''(x) < 0$, $p(0) = 0$ and $p(\infty) = 1$[11].

Suppose $\theta = 0$. For simplicity, assume the firm is small in the sense there is no strategic interaction in R&D. Thus, the dynamic returns to the knowledge k are:

[10] Compared to equation (2.3), where Z is divided by the term $1 - p(1-q)$, this term is the probability that the line of research will be terminated *in firm i's perspective*, which occurs when firm i fails to win in a development stage.

[11] Decreasing returns to scale in the production of knowledge are necessary to ensure an interior solution

$$W(\theta = 0) = \frac{vp(x)(1-q) - x}{1 - p(x)(1-q)} \tag{2.8}$$

The firm maximizes equation (2.8) with respect to its R&D expenditures, x, which yields the following first order condition:

$$p'(x|\theta = 0) = \frac{1}{(1-q)(W(\theta = 0) + v)} \tag{2.9}$$

Let $x^*(\theta = 0)$ solve equation (2.9). Thus, the optimality condition equates the marginal benefit from R&D (the increase in the probability of a discovery that is achieved by a marginal increase in the R&D spending, $p'(x)$), adjusted by the probability that the firm will be the sole winner in the research race (while taking into account the one shot pay-off, v, and the total pay-off of winning the race, $W(\theta = 0)$), to the marginal cost of R&D, which is assumed to be 1. An increase in q reduces the probability of winning the development stage and, consequently, reduces the R&D expenditures of the firm (note that $W(\theta = 0)$ is a decreasing function of q). On the other hand, an increase in the one shot payoff, v, encourages the firm to innovate more (note that $W(\theta = 0)$ is an increasing function of v). More importantly, a rise in the dynamic rent, given by $W(\theta = 0)$, increases the innovation efforts of the firm.

Now consider the case where $\theta = \infty$. Dynamic returns to knowledge k are expressed as:

$$W(\theta = \infty) = \frac{vp(x)(1-q) - x}{(1 - p(x))(1-q)} \tag{2.10}$$

---

for x. The second order condition is:

$$\frac{\partial W^2(\cdot)}{\partial^2 x} = p''(x) - \frac{1}{(1-q)}\left[-\frac{\frac{\partial W(\cdot)}{\partial x}}{(W(\cdot) + v)^2}\right] \leq 0$$

Define x^* as the optimal R&D decision, thus:

$$\frac{\partial W^2(\cdot)}{\partial^2 x}|_{x=x^*} = p''(x)$$

For x^* to be a maximum we require $p''(x) < 0$.

The first order condition for x is:

$$p'(x|\theta=\infty)=\frac{1}{(1-q)(W(\theta=\infty)+v)} \tag{2.11}$$

Let $x^*(\theta=\infty)$ solve equation (2.11).

Proposition 2.1. *The firm innovates more in case $\theta=\infty$, compared to the case $\theta=0$.*

To prove proposition 2.1, it is enough to show that $W(\theta=\infty, x^*(\theta=\infty)) > W(\theta=0, x^*(\theta=0))$, since $p''(x)<0$. Suppose that $W(\theta=0, x^*(\theta=0)) > W(\theta=\infty, x^*(\theta=\infty))$. This inequality cannot hold, since I have shown above that $W(\theta=\infty, x) > W(\theta=0, x)$. Thus, $W(\theta=\infty, x^*(\theta=0)) > W(\theta=\infty, x^*(\theta=\infty))$, which is a contradiction of $x^*(\theta=\infty)$ being the optimal R&D investment when $\theta=\infty$. Hence, it must be that $W(\theta=\infty, x^*(\theta=\infty)) > W(\theta=0, x^*(\theta=0))$, which implies that $p'(x|\theta=\infty) < p'(x|\theta=0)$ and $x^*(\theta=\infty) > x^*(\theta=0)$. The case where $W(\theta=\infty, x^*(\theta=\infty)) = W(\theta=0, x^*(\theta=0))$ cannot hold from exactly the same argument[12].

The simple intuition behind this proposition is that the firm is willing to invest more in R&D, when private returns are higher.

In conclusion, a higher θ should intensify technological internalization and, therefore, raise private returns. In case firms are aware of the effect of θ on private returns, R&D expenditures should rise in θ as well.

The rest of the paper empirically supports these theoretical predications. In the econometric section, dynamic private returns (W_i) are specified as $\frac{\partial V_i}{\partial K_i}$, where V_i is the market value of the inventing firm and K_i is its knowledge stock (which is approximated by current and past stream of own R&D expenditures)[13]. Private returns are specified as a function of θ, $\frac{\partial V_i}{\partial K_i}=\Theta(\theta)$, and θ is empirically identified by measures of Internalized and Externalized spillovers.

[12] Appendix A.1 shows that proposition 2.1 can be generalized for every θ.

[13] It is convenient to model private returns as the effect of the knowledge of the inventing firm on its market value, since markets are forward looking and should incorporate the dynamic consideration of technological internalization developed in this paper.

3. Methodology

This section discusses the conceptual issues that underpin the empirical framework of measuring technological internalization. I start by presenting how the technological contribution of an invention is measured. Then, spillovers are defined as the external exploitation of the technological contribution of the invention. Finally, it is shown how it is determined whether spillovers feed back into the inventing firm to generate technological internalization.

3.1. Identifying the technological contribution of an invention

I propose measuring the technological contribution of an invention in two dimensions. The first is the number of *lines of research* the invention originates and the second is the 'quality' of these lines of research. A line of research is defined as *a sequence of inventions, where every invention is a follow-up development of its immediate ancestor.* This sequence of inventions is required to be unique over a given time period, i.e., not to be fully contained in a longer sequence of inventions. Define the first invention in the line of research as an *originating invention.* A line of research is assumed to be of a higher 'quality', if the number of subsequent developments of the originating invention along the line of research is higher.

More formally, the technological contribution of invention i, TC_i, is computed as the 'quality'-weighted count of the lines of research invention i originates, as following[14]:

$$TC_i = \sum_{k \in K_i} LR_k \times Q_k \tag{3.1}$$

Where, K_i is the set of lines of research originated in invention i, k indexes lines of research in this set, LR_k is a dummy that receives the value 1 for line of research k and zero otherwise, and Q_k is the 'quality' of line of research k, as measured by the number of

[14] Belenzon (2005) shows that this method of measuring technological contribution is equivalent to an alternative approach of counting the number of offspring inventions and weighing each one by the number of direct citations received.

inventions the line of research includes[15].

Applying this formulation to the diffusion patterns in figure 1 yields:

$$TC_A^1 = (1 \times 3) = 3 \tag{3.2}$$

Where, TC_A^1 is the technological contribution of invention A under pattern 1. The term 1 in the brackets represents the singleton line of research $A \rightarrow B \rightarrow C \rightarrow D$ that is adjusted by its 'quality', which is 3 (since it includes three subsequent developments of invention A: B, C and D).

Similarly, the technological contribution of invention A under diffusion pattern 2, TC_A^2, is:

$$TC_A^2 = (1 \times 2) + (1 \times 2) = 4 \tag{3.3}$$

The term 1 in the first brackets represents the line of research $A \rightarrow B \rightarrow C$ that is adjusted by its 'quality', which is 2 (since it includes two subsequent developments of invention A: B and C). The term 1 in the second brackets represents the line of research $A \rightarrow B \rightarrow D$ that is adjusted by its 'quality', which is 2 as well (since it includes two subsequent developments of invention A: B and D).

From this is concluded that the technological contribution of invention A under diffusion pattern 2 is greater than its technological contribution under diffusion pattern 1 (intuitively, under both patterns of diffusion the number of subsequent developments is equal. However, there are more research opportunities under pattern 2, as indicated by the number of lines of research).

[15] Simply counting the number of inventions along a line of research may be an overestimate of the technological contribution of the originating invention. A subsequent invention which is a high generation of development of the originating invention is more likely to have benefited from other prior subsequent inventions along the line of research. Thus, I always discount every generation by a discount factor of δ per generation (which is assumed to be 15 percent), thus, $Q_k = \sum_{j=1}^{J} \delta^{j-1}$, where, J is the number of offspring inventions in line of research k. Since the choice of the discount factor is arbitrary, other values of δ are experimented with as robustness tests.

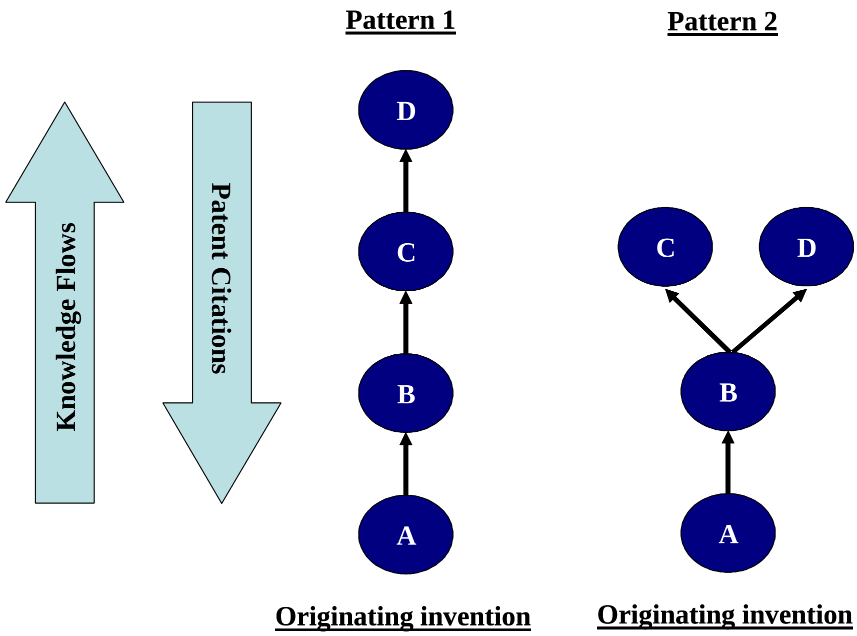


Figure 2: Technological contribution

Figure 2: *Circles in this figure represent inventions and arrows represent the direction of knowledge flow. Pattern 1 illustrates a singleton path of knowledge flow, which is* $A \rightarrow B \rightarrow C \rightarrow D$*, while diffusion pattern 2 illustrates two unique paths of knowledge flows, which are* $A \rightarrow B \rightarrow C$ *and* $A \rightarrow B \rightarrow D$*. Determining the technological contribution of invention* A *under the two diffusion patterns requires weighing these lines of research by their 'quality', by measuring their length in terms of the number of inventions they include.*

3.2. Measuring spillovers

Spillovers are defined as the external exploitation of the technological contribution of an invention, where *external* refers to the set of firms that are different from the inventing firm. Following this definition, spillovers are measured as the number of external inventions along the lines of research the originating invention inspires.

For illustration, it is useful to examine a slightly more complicated diffusion pattern, as shown in figure 2. Capital letters represent inventions, where arrows represent the direction of knowledge flow. This figure plots the diffusion pattern of the originating invention A, where the offspring inventions are B, C, D, E, F, G, H, I and J. To complete the presentation, the shape of each capital letter represents a different firm, i.e., a circle firm (the inventing firm), a triangle firm and a square firm.

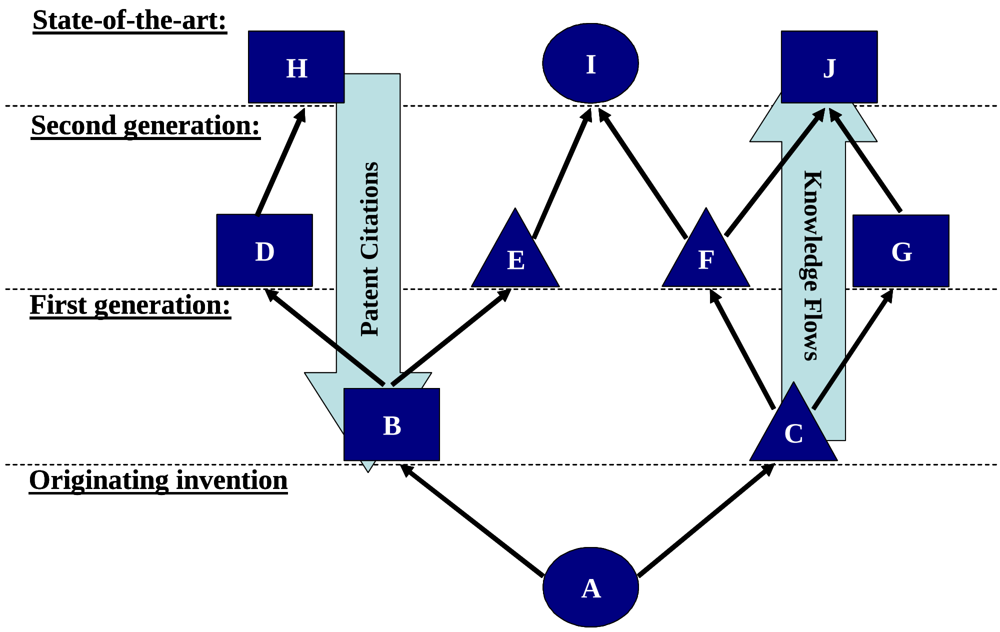


Figure 3: Measuring spillovers

Figure 3: *This figure illustrates the diffusion pattern of the originating invention A. Inventions are represented by a capital letter, while the firm that owns the inventions is represented by a shape (e.g., the inventing firm is the circle, since it owns the originating invention A). I define the spillovers created by invention A, given this diffusion pattern, as the number of inventions that are owned by the square and triangle firms (all the firms in the figure which are different from the inventing firm) along the lines of research invention A originates.*

Following the methodology presented above, in order to measure the technological contribution of invention A, we need to identify the lines of research invention A originates and weigh them by their 'quality'. Since a line of research is defined as a singleton sequence of subsequent developments of the originating knowledge, there are five such lines of research: $A \rightarrow B \rightarrow D \rightarrow H$, $A \rightarrow B \rightarrow E \rightarrow I$, $A \rightarrow C \rightarrow F \rightarrow I$, $A \rightarrow C \rightarrow F \rightarrow J$ and $A \rightarrow C \rightarrow G \rightarrow J$. The technological contribution of invention A following equation (3.1) is given by:

$$TC_A = (1 \times 3) + (1 \times 3) + (1 \times 3) + (1 \times 3) + (1 \times 3) = 15 \tag{3.4}$$

Since spillovers are defined as the external inventions that compose the lines of research an invention originates, they are formulated as:

$$Spillovers_i = \sum_{k \in K_i} LR_k \times S_k \tag{3.5}$$

Where, i is an originating invention, K_i is the set of lines of research invention i originates, k indexes lines of research in this set, LR_k is a dummy that receives the value 1 for line of research k and zero otherwise and S_k is the number of external inventions included in line of research k. Following this formulation, the spillovers created by invention A are:

$$Spillovers_A = (1 \times 3) + (1 \times 2) + (1 \times 2) + (1 \times 3) + (1 \times 3) = 13 \tag{3.6}$$

Where, the second and third terms, (1×2) and (1×2), correspond to the fact that invention I is owned by the inventing firm. Thus, invention I is excluded from the spillovers measure for invention A (the spillovers along lines of research $A \rightarrow B \rightarrow E \rightarrow I$ and $A \rightarrow C \rightarrow F \rightarrow I$ are based only on inventions B, E, C and F)[16].

[16] In some patterns of diffusion, the first subsequent development of the originating knowledge is done by the inventing firm (which is identified as a *self-citation*). Hence, knowledge does not immediately spread to other inventors. In this case, the 'in-house' subsequent development is not measured as spillovers (where spillovers along such lines of research occur only if in a future generation of development knowledge leaves the boundaries of the inventing firm).

Finally, I aim at distinguishing between two types of spillovers: spillovers that contribute to the dynamic research of the inventing firm and spillovers that do not.

3.3. Internalized and Externalized lines of research

Two types of lines of research are identified: the first type is lines of research where the originating knowledge leaves the inventing firm and returns to this firm after having been further developed by other firms. The second type is lines of research where the originating knowledge leaves the inventing firm and does not return. Spillovers along the former type are *internalized* in the dynamic research of the inventing firm and, therefore, these lines of research are defined as *Internalized lines of research*. However, spillovers along the latter type do not contribute to the dynamic research of the inventing firm, therefore, these lines of research are defined as *Externalized lines of research*.

Hence, the spillovers of an invention can be written as:

$$Spillovers_i = \sum_{j \in Internalized_i} LR_j \times S_j + \sum_{t \in Externalized_i} LR_t \times S_t \tag{3.7}$$

Where i denotes an originating invention, $Internalized_i$ is the set of Internalized lines of research originated in invention i, $Externalized_i$ is the set of Externalized lines of research originated in invention i, j indexes lines of research in the $Internalized_i$ set and t indexes lines of research in the $Externalized_i$ set. I define the first term in the right-hand-side of equation (3.7) as $IntSpill_i$ and the second term in the right-hand-side of equation (3.7) as $ExtSpill_i$. Thus, equation (3.7) becomes:

$$Spillovers_i = IntSpill_i + ExtSpill_i \tag{3.8}$$

In addition, $IntShare_i$ is defined as the ratio between $IntSpill_i$ and $Spillovers_i$.

To illustrate this decomposition, it is useful to refer back to figure 2. Out of the five lines of research that invention A originates, two are Internalized and three are Externalized. The set $Internalized_A$ is:

$$Internalized_A = \{A \to B \to E \to I, A \to C \to F \to I\}$$

Similarly, the set $Externalized_A$ is:

$$Externalized_A = \{A \to B \to D \to H, A \to C \to F \to J, A \to C \to G \to J\}$$

Given this decomposition, $IntSpill_A = (1 \times 2) + (1 \times 2) = 4$ (two external inventions in the first line of research and two external inventions in the second line of research in the $Internalized_A$ set).

Similarly, $ExtSpill_A = (1 \times 3) + (1 \times 3) + (1 \times 3) = 9$ (three external inventions in each of the three lines of research in the $Externalized_A$ set).

3.4. Empirical methodology

Inventions are empirically identified as patents and knowledge flows as citations (where knowledge flows from the cited patent to the citing patent). Patents and citations data contain significant noise and bias[17]. Nonetheless, these data also offer unique information on the diffusion pattern of knowledge and sequential innovation, which I believe to be extremely useful for exploring the ideas developed in this paper.

Hence, inventions in figures 2 and 3 are empirically identified as patents, whereas arrows are empirically identified as citations. For example, an arrow from invention A to invention B in figures 2 and 3 reflects the fact that patent B cites patent A. The task I am facing is to effectively draw figure 2 for the set of originating inventions[18].

A unique line of research is empirically identified as a *singleton sequence of citations* (where, each patent cites its direct ancestor). As discussed above, a sequence of citations is defined as singleton, if it is not fully contained in a longer sequence of citations

[17] See, for example, Trajtenberg (1990) for the potential bias in patents as indicators for innovation output, and Trajtenberg, Jaffe and Fogarty (2001) for a study on the noise component in citations as indicators for knowledge flows.

[18] The design of this set is explained below.

for the given time period being explored. After extracting the lines of research for the sample of originating patents, each line of research is classified as either Internalized or Externalized[19].

The period for which lines of research are constructed is restricted to 15 years after the grant year of the originating patent. For example, for a patent that was granted in 1975, the youngest patents in all the lines of research it originates cannot be granted after 1990. Further, citations along a line of research are added as long as the line of research has not already been classified as Internalized[20]. Thus, this methodology extracts all the unique trajectories where knowledge had left the boundaries of its inventor and returned to these boundaries in a time period of 15 years after the knowledge had been created[21], as well as all the unique trajectories where knowledge had left the boundaries of the inventing firm and did not return to these boundaries in the same time period[22].

Since this paper exploits the firm-level variation in technological internalization, $IntSpill$, $ExtSpill$ and $IntShare$ are aggregated to the firm level by the taking their mean over the set of originating patents held by the inventing firms. For ease of notations, these variables are not re-labeled. Hereafter, they refer to the firm-level aggregates.

[19] The reader who is familiar with the economics of patents literature can find the definition of an Internalized line of research similar to a self-citation. A self-citation is the case where a firm develops its prior knowledge directly (the first citation the patent receives is from the inventing firm). An Internalized line of research is the case where the firm *indirectly* develops its prior knowledge, after it has been developed by other firms. Thus, an Internalized line of research is a unique *indirect self-citation*, which I associate with a higher appropriability, as the existing literature does with self-citations (e.g., Hall, Jaffe and Trajtenberg (2005)).

[20] E.g., consider the Internalized line of research $A \rightarrow B \rightarrow E \rightarrow I$ that is presented in figure 3. Assume that patent I is cited by patent K, such that this line of research becomes $A \rightarrow B \rightarrow E \rightarrow I \rightarrow K$. The imposed restriction implies that only the line of research $A \rightarrow B \rightarrow E \rightarrow I$ will be extracted for the originating patent A.

[21] Since I refer to the grant year of the patent and not to its application year, the creation date of the patented knowledge is actually earlier. However, my algorithm builds on the fact that a citing patent cannot be cited before it cites. This crucial feature of the data can be exploited only by referring to the grant year of the patent (see Belenzon (2005) for detail on the algorithm).

[22] It is important to note that this methodology incorporates the case where knowledge is first developed sequentially 'in-house' by the inventing firm (i.e., self-citations). In numerous cases the inventing firm develops the first follow-up inventions of the originating knowledge. In such lines of research knowledge leaves the boundaries of the inventing firm via a higher order generation of citation. These lines of research are classified as Internalized or Externalized following the same criterion described above.

4. Data

4.1. Diffusion data

Patents and citations data are taken from the U.S. Patent and Trademark Office from the NBER archive. The sample of patents and citations includes about 1.7 million citations and 600,000 patents (which can appear as offspring inventions)[23]. The set of originating patents includes all patents granted between 1969 and 1980[24] that received at least one citation during 1975-1995 (recall that all the direct and indirect citations the originating patent receives are extracted for a period of 15 years since its grant year). These patents must be held by the sample of firms for which complete accounting data are available for the period 1980-2001 (*IntSpill* and *ExtSpill* are computed for patents from a pre-estimation period. It is assumed that the pattern of diffusion of this set of inventions is a time invariant characteristic of the firm[25]). This set of originating patents includes 104,694 patents[26] (see appendix A.2.1 for more detail). Detail on the algorithm developed to construct the diffusion data is provided in Belenzon (2005).

Table 1 describes the variation of lines of research across technology sectors and time. The largest number of lines of research per citation received by an originating patent is in the "Electrical and Electronics" sector. This may indicate a high technological

[23] The set of citing patents includes all patents held by the US Compustat firms that were matched to the USPTO by Hall, Jaffe and Trajtenberg (2001), and made at least one citation. This set includes about 30 percent of all citing patents in the USPTO (and 50 percent of US citing patents).

[24] The year 1969 is the earliest year for which there is citations information for the patents held by the firms in the sample. Also, in practice I could extract the diffusion pattern of patents that were granted up to 1985, since the citations data goes up to 1999. However, there is a huge spike in the number of citations in 1995 (see figure A3), where the number of citations rises by around 800,000 in the period 1995-1999. In addition to the feasibly of extracting sequences of citations from these huge data, there is also a concern that the explosion in citations in this period is not associated with stronger learning and sequential innovation, but with changes in the patenting behavior of firms, which could contaminate the results.

[25] There is a trade-off between constructing time varying diffusion variables and correctly measuring *IntSpill* and *ExtSpill*. In case the 15 years horizon for which sequences of citations are extracted for every patent is reduced, more periods for the construction of *IntSpill* and *ExtSpill* would be observed. However, when analyzing a short diffusion period the lines of research are more likely to be Externalized, as knowledge has less time to return to the inventing firm.

[26] This set includes 45 percent of all cited patents between 1969 and 1980 that are held by US Compustat firms that were matched to the USPTO by Hall, Jaffe and Trajtenberg (2001).

complexity in this sector, where complexity refers to the various distinct ways along which knowledge can be sequentially developed[27]. About 7.6 percent of the lines of research are Internalized. This share appears to be rather stable over time, with an exception in "Drugs and Medicals". In the period 1978-1980 there is a large drop in the share of Internalized lines of research in this sector, which may be associated with the Biomed revolution that took place at the end of the 70's. I plan to investigate this separately in a future research.

[**Table 1 about here**]

4.2. Accounting data

The accounting data (sales, R&D, capital, etc.) and market value data are taken from US Compustat for the period 1980-2001 and are merged to the U.S. Patent and Trademark Office data from the NBER archive[28]. Only firms that were cited during the diffusion period were included in the sample, leaving an unbalanced panel of about 500 firms in the period 1980-2001 and a total of 9,454 observations. The final sample includes the largest patenting firms in the US. The average number of years firms are active in the sample is 18.5 (and the median is 21). I find it important to focus the analysis on long surviving firms, due to the interest of studying the effect of technological internalization on firms that wish to remain at the frontier of the technology they originate. Otherwise, Externalized spillovers could capture exit of firms (since firms that exit will have only Externalized spillovers from the date they exit onwards), which will change the interpretation of technological internalization[29]. Appendix A.2 provides detail on the construction of this sample.

Table 2 summarizes the descriptive statistics for $IntSpill$, $ExtSpill$ and $InShare$ as well as for the main accounting variables. The correlation between $IntSpill$ and $ExtSpill$

[27]For example, technology field 438 (Semiconductor Device Manufacturing: Process) in the "Electrical and Electronics" sector has on average 112.3 lines of research per citation received by an originating patent. On the other hand, technology field 139 (Textiles: Weaving) in the "Chemicals" sector has only of 5.1 such lines of research.

[28]The matching between assignee names and Compustat firms is taken from Hall, Jaffe and Trajtenberg (2001).

[29]It would be interesting to analyze the effect of technological internalization on exit, however, this is not studied in this paper.

is 0.365 (which implies that firms that create more Internalized spillovers also create more Externalized spillovers). About 40 percent of firms do not create Internalized spillovers at all, whereas all firms create Externalized spillovers.

[Table 2 about here]

5. Estimation

This section presents the estimation results of the effects of $IntSpill$ and $ExtSpill$ on the market value and R&D expenditures of the firm. I start by showing that these variables matter for the market valuation of the R&D stock of the firm. Findings that technological internalization matters for market value may imply that it also matters for the R&D decisions of firms. I test this hypothesis by estimating a R&D expenditures equation.

5.1. Market value equation

In order to estimate the effect of technological internalization on private returns, a simple version of the value function approach proposed by Griliches (1981)[30] is adopted. The market value of firm i at period t, V_{it}, takes the following form:

$$V_{it} = \kappa_{it} \left(A_{it} + \gamma K_{it} \right) \tag{5.1}$$

Where, A_{it} denotes physical assets, K_{it} is the R&D stock (representing knowledge stock), γ is the shadow price of the R&D stock (higher values of γ indicate that the market valuation of the knowledge stock relative to physical stock rises)[31]. The parameter γ captures the private returns to innovation, which are defined as the change in market value as a response to a change in the R&D stock of the firm. γ is modeled as a linear function of $IntSpill$ and $ExtSpill$[32]:

$$\gamma = \gamma_0 + \gamma_1 \left(IntSpill_i \right) + \gamma_2 \left(ExtSpill_i \right) \tag{5.2}$$

[30] See also Jaffe (1986), Hall et al (2005) or Lanjouw and Schankerman (2004).

[31] A constant returns in the market value function has been assumed, consistently with previous studies.

[32] Specifications where $IntShare$ is included instead of $IntSpill$ and $ExtSpill$ are also reported.

I expect γ_1 to be positive and γ_2 to be negative (the theoretical predication is that private returns rise with θ, empirically captured by a higher $IntSpill$ and a lower $ExtSpill$).

Taking logarithms and dividing by A_{it}, the left-hand-side of equation (5.1) becomes the traditional Tobin's average Q, where its deviation from unity depends on the ratio between the R&D stock to the tangible stock $\left(\frac{K}{A}\right)$, $IntSpill$, $ExtSpill$ and κ_{it}, as following:

$$\log\left(\frac{V_{it}}{A_{it}}\right) = \log\kappa_{it} + \log\left(1+\gamma\frac{K_{it}}{A_{it}}\right) \tag{5.3}$$

Finally, κ_{it} is specified as:

$$\log\kappa_{it} = Z_{it}'\beta_0 + \beta_1\log\left(1+IntSpill_i\right) + \beta_2\log\left(1+ExtSpill_i\right) + \tau_t + \eta_i + \epsilon_{it} \tag{5.4}$$

Where, Z_{it} is a vector of controls (such as industry and technology dummies, sales, patents stock, etc.), τ_t is a complete set of time dummies, η_i is the firm fixed-effect, which is discussed later in this section, and ϵ_{it} is an idiosyncratic error term. The linear terms of $IntSpill$ and $ExtSpill$ are included in the specification mainly as controls for their interaction with the R&D stock. Since $IntSpill$ has many zero values, a dummy for $IntSpill$ equals zero is always included.

Thus, the following equation is estimated by non-linear least squares (where standard errors are clustered by firms):

$$\begin{aligned}\log\left(\frac{V_{it}}{A_{it}}\right) &= Z_{it}'\beta_0 + \beta_1\log\left(1+IntSpill_i\right) + \beta_2\log\left(1+ExtSpill_i\right) \\ &\quad +\log\left(1+\left(\gamma_0+\gamma_1\left(IntSpill_i\right)+\gamma_2\left(ExtSpill_i\right)\right)\frac{K_{it}}{A_{it}}\right) + \tau_t + \eta_i + \epsilon_{it}\end{aligned} \tag{5.5}$$

5.2. R&D equation

A R&D equation is estimated in order to test whether $IntSpill$ and $ExtSpill$ affect the R&D decision of firms, as predicted by proposition 2.1 (intuitively, in case technological internalization raises private returns, we should find firms with higher technological internalization, on average, investing more in R&D). I estimate the firm fixed-effects in the R&D equation (a complete set of firm dummies), and project the estimated firm fixed-effects on $IntSpill$ and $ExtSpill$.

The R&D equation that is estimated is:

$$\log R\&D_{it} = \alpha_i + X_{it}'\beta + \epsilon_{it} \tag{5.6}$$

Where, $R\&D_{it}$ is the R&D expenditures of firm i in period t, α_i is the firm fixed-effect, X_{it} is a vector of controls for sales and patents variables and ϵ_{it} is an idiosyncratic error term. The sales variables (current and lagged) aim to capture demand shocks that may affect R&D incentives. The patents stock (weighed by citations) can affect the R&D decision of the firm in various ways. One possibility is that patents capture the intellectual property protection the firm faces, so that a larger patent portfolio raises the incentive to perform R&D. I also estimate a dynamic specification by adding $R\&D_{it-1}$ in the right-hand-side of equation (5.6).

Based on the estimates obtained from the R&D equation, $\widehat{\alpha}_i$ is projected in the second stage on $IntSpill$ and $ExtSpill$:

$$\widehat{\alpha}_i = \delta_1 \log\left(1 + IntSpill_i\right) + \delta_2 \left(1 + ExtSpill_i\right) + \delta_3 \overline{X}_i' + \nu_i \tag{5.7}$$

Where, $\overline{X}_i$ is the mean of X_{it} over the estimation period (1980-2001) and ν_i is the error term. δ_1 is expected to be positive and δ_2 is expected to be negative[33]. Since $IntSpill$ has many zero values, a dummy for $IntSpill$ equals zero is always included.

[33] In addition, the estimation results of including only $IntShare$ are reported. $IntShare$ is expected to have a positive effect on $\widehat{\alpha}_i$.

5.3. Preliminary issues for the market value estimation

5.3.1. Dealing with cross-industry variation

Firms in the sample are located in different industries. These industries may vary in private returns and technological internalization. Hence, a pooling estimation across industries may capture industry variation in private returns, via *IntSpill* and *ExtSpill*. I cope with this concern in various ways. In all the specifications reported below, complete sets of two-digit industry dummies and main technology sector indicators are included[34]. The technology indicators are the share of the firm's patents in each of the five main technology sectors. Moreover, small sample evidences are presented for three specific heterogenous industries.

Table 3 looks at the variation of the diffusion variables across four levels of industry aggregation. The analysis of the variation of the diffusion variables shows that the main variation comes from within industries, mainly for *IntSpill* and *ExtSpill* (*IntShare* appears to be more sensitive to industry effects, however, about 50 percent of its variation is still evident within four-digit industry breakdown). This finding is encouraging, since it is more likely that the source of variation in technological internalization is not strongly associated with industry location.

[Table 3 about here]

5.3.2. Endogeniety

The use of firm level accounting data may lead to the classical endogeniety bias in the R&D stock. A higher market value can, indeed, be the result of conducting more R&D, however, the ability to devote more resources to R&D can reflect a higher market value that provides more finance to the innovative activity. Moreover, demand or supply shocks can simultaneously raise the R&D expenditures and the market value of the firm. In order to mitigate this potential bias, a complete set of year dummies is included, aiming at

[34] The estimation results of a linearized version of equation (5.5) with a complete set of four-digit SIC dummies are also reported.

capturing transitory shocks[35]. Further, my main interest is to recover the effects of *IntSpill* and *ExtSpill*, which are less sensitive to the endogeniety of the accounting variables.

A more serious endogeniety bias may be associated with *IntSpill* and *ExtSpill*, as there is an overlap in the period used for their construction and the estimation period[36]. In order to test the sensitivity of the findings, I experiment with different time periods for the construction of *IntSpill* and *ExtSpill*, so as to reduce the overlap with the estimation period[37]. The pattern of results is robust to the different time periods.

5.3.3. Firm fixed-effects

I do not control for firm fixed-effects in the market value equation by including a complete set of firm dummies from two main reasons: first, under the assumption of efficient markets, changes in market value should not be predicted (especially by common observable characteristics, such as R&D stock)[38]. Second, *IntSpill* and *ExtSpill* are time-invariant. Thus, in the presence of firm dummies, the only way to identify their effect is via the variation in their interaction with the R&D stock. Since the R&D stock is rather persistent over time within firms, in practice, there is no significant effect of *IntSpill* and *ExtSpill* when a complete set of firm dummies is included.

Therefore, I control for firm fixed-effects by adopting the "mean scaling" approach developed by Blundell, Griffith and Van Reenen (1999). Their method assumes that computing the mean of Tobin's Q in a long enough pre-estimation period can be used as an initial condition to proxy for unobserved heterogeneity, if the first moment is stationary. In order to amplify the effectiveness of this method and test its robustness, I also include the pre-estimation means of other firm-level variables, such as sales, industry sales, employees, R&D stock, citations-weighted patents stock and citations stock[39]. The pre-sample means

[35] I also experiment with lagging the R&D stock by one period, which yields similar results.

[36] The estimation period is 1980-2001, whereas the diffusion variables are constructed for the period 1969-1995.

[37] *IntSpill* and *ExtSpill* are also constructed for the periods: 1969-1990, 1969-1985 and 1969-1980.

[38] Hall, Jaffe and Trajtenberg (2005) reach a similar conclusion.

[39] Including the mean of additional right-hand side variable is also important since they are computed for a similar period used for the construction of *IntSpill* and *ExtSpill*. For example, in case technological

of the accounting variables are constructed from US Compustat for the period 1970-1979 for the 476 firms in the estimation sample.

5.4. Estimation results for Tobin's Q

All the Tobin's Q specifications include a complete set of two-digit industry dummies (79 dummy variables), a set of indicators for the share of patents the firm has in each of the five main technology sectors, a complete set of year dummies (21 dummy variables), a dummy variable that receives the value one if the R&D stock of the firm is zero and a dummy variable that receives the value one if *IntSpill* is zero.

Table 4 reports the estimation results of equation (5.5). Column 1 includes the linear term of R&D over assets and its interaction terms with *IntSpill* and *ExtSpill*. The coefficient on the linear term of R&D over assets (γ_0) is positive and significant (0.280 with a standard error of 0.079). The coefficient on the interaction term of *IntSpill* with R&D over assets (γ_1) is positive and significant (0.208 with a standard error of 0.095), while the coefficient on the interaction term of *ExtSpill* with R&D over assets (γ_2) is negative and significant (-0.011 with a standard error of 0.003). These findings support the expectation that private returns rise with technological internalization.

Given these estimates, the elasticity of market value with respect to the R&D stock, evaluated at the sample mean, is 0.103[40]. This implies that an additional one dollar spent on R&D raises market value by 0.49 dollar (referred to as private returns). A one standard deviation increase in *IntSpill* raises private returns to 0.63 dollar (thus, a 30 percent increase), whereas a one standard deviation increase in *ExtSpill* lowers private returns to 0.44 dollar (thus, a 10 percent decrease)[41].

---

internalization is higher when the firm has more originating patents, including only the patents stock in the estimation period is not sufficient.

[40] The estimated elasticity is lower from that reported in previous studies. For example, Bloom, Schankerman and Van Reenen (2005) report an elasticity of 0.24, using a similar estimation sample without industry or technology effects.

[41] The estimated effects of *IntSpill* and *ExtSpill* on private returns are underestimated, as it is assumed that a change in either measures is independent from the other. For example, it is likely that an increase in *IntSpill* will reduce *ExtSpill* as well (thus, a line of research becomes Internalized instead of being Externalized). This indicates that private returns will rise as a result of the increase in *IntSpill* and also

In column 2, the pre-sample means are added[42]. The coefficient on the linear term of R&D stock over assets halves (from 0.280 to 0.145) and remains significant. The coefficient on the interaction term of *IntSpill* with the R&D stock over assets drops from 0.208 to 0.096, but it remains significant, while the coefficient on the interaction term between *ExtSpill* and the R&D stock over assets drops in absolute value (from -0.011 to -0.005) and remains significant as well[43].

Based on the estimates from column 2, the elasticity of market value with respect to the R&D stock, evaluated at the sample mean, is 0.056 (compared to 0.103 without pre-sample means). An additional one dollar spent on R&D raises market value by 0.26 dollar (compared to 0.49 dollar without the pre-sample means). A one standard deviation increase in *IntSpill* raises these private returns to 0.34 dollar, while a one standard deviation increase in *ExtSpill* lowers private returns to 0.24 dollar.

In column 3, *IntSpill* and *ExtSpill* are added linearly. The same pattern of results holds. The main change is a drop in the coefficient on the interaction term of *IntSpill* (from 0.096 to 0.059), which remains significant. The coefficient on the linear term of *IntSpill* is positive and significant, while the coefficient on the linear term of *ExtSpill* is negative and significant, both as expected. Importantly, the positive effect of *IntSpill* and the negative effect of *ExtSpill* are identified linearly and through their interaction with the R&D stock over assets.

In columns 4 and 5, the sales of the firm, the aggregate sales in the industry the firm operates in and the growth in the sales of the firm are added (see appendix A.2 for detail on their construction). The same pattern of results with respect to *IntSpill* and *ExtSpill* (linear and interacted terms) remains. The effects of Sales and Sales Growth are positive and significant and the effect of Industry Sales is negative and significant.

---

as a result of a decrease in *ExtSpill*.

[42] The set of pre-sample means is jointly significant with a *p-value*<0.001.

[43] I also interact the pre-sample mean of Tobin's Q with R&D over assets, in order to test the robustness of the interaction terms of *IntSpill* and *ExtSpill*. The coefficient on the interacted term of *IntSpill* rises to 0.112 with a standard error of 0.019 and the coefficient on the interacted term of *ExtSpill* is -0.006 with a standard error of 0.001.

Table A2 reports the same estimations for $IntShare$. The effect of $IntShare$ on market value is always positive and significant.

[Table 4 about here]

Table 5 summarizes the quantitative effects of *IntSpill*, *ExtSpill* and $IntShare$ on private returns in terms of percentage point changes in market value as a response to a 1 dollar increase in R&D expenditures, evaluated at the mean (the columns in table 5 correspond to the same columns in tables 3 and A2). Including pre-sample means does not change the effect of *IntSpill* and *ExtSpill* (comparing column 1 to column 2), however, it raises the effect of $IntShare$. Adding the linear terms *IntSpill*, *ExtSpill* and $IntShare$ (column 3) substantially lowers the effect of the interaction terms.

[Table 5 about here]

5.5. Estimation results for the R&D equation

Table 6 reports the estimation results for equations (5.6) and (5.7)[44]. I explore two specifications: first, only *IntSpill* and *ExtSpill* are included, aiming at identifying the extent cross-firm variation in the time-invariant component in the R&D decision is attributed to technological internalization. Second, the means of the first-stage variable in the period 1980-2001 are added.

Column1 reports the results from the first-stage estimation for the static specification. Current and lagged sales have a positive effect on R&D expenditures, where the effect of industry sales is positive and significant only for the current term. A possible interpretation of the positive effect of sales on R&D expenditures is transitory shocks which raise the sales (both of the firm and its competitors) and the incentive to innovate. The citations-weighted patents stock has a positive and significant effect on R&D expenditures. This is consistent with the expectation that the citations-weighed patent stock captures higher appropriation through stronger intellectual property protection, which raises the incentive to perform R&D.

[44] In the dynamic specifications, the *long-run* fixed-effect is computed as $\frac{\widehat{\alpha}}{1-0.568}$, where 0.568 is the coefficient on the lag of R&D.

Column 2 reports the estimation of a dynamic specification of the R&D equation. A similar pattern of results holds, with the exception that the coefficient on lagged sales is negative and significant.

In column 3, the firm fixed-effects obtained from the first-stage static estimation are projected on $IntSpill$ and $ExtSpill$. The effect of $IntSpill$ is positive and significant, whereas the effect of $ExtSpill$ is negative and significant, both as expected. The R^2 is 0.39, which indicates that about forty percent of the between-firm variation in the R&D decision can be explained by technological internalization. In column 4, the means over the estimation period of the first-stage variables are added. The coefficients on $IntSpill$ and $ExtSpill$ drop, however, they keep their signs and remain significant.

Columns 5 and 6 report the equivalent estimation results for the fixed–effects from the dynamic specification. The same pattern of results holds.

Finally, columns 7 to 10 report the same estimation for $IntShare$, which show similar results.

Based on the static estimations, a one standard deviation increase in $IntSpill$ raises R&D expenditures by 33 percent, whereas a one standard deviation increase in $ExtSpill$ lowers R&D expenditures by 5 percent. Further, a one standard deviation increase in $IntShare$ raises R&D expenditures by 38 percent. Thus, the quantitative effects of Internalized spillovers are large, whereas the effect of Externalized spillovers is much lower. This implies that what matters most for the R&D decision of the inventing firm is the spillovers that feed back into its research, rather spillovers that do not[45].

[Table 6 about here]

[45] I also estimate the effect of $IntSpill$, $ExtSpill$ and $IntShare$ on R&D expenditures in a one stage, where these variables are added to the first stage estimation reported above. For the static estimation, the coefficients on $log(1+IntSpill)$ and $IntShare$ are positive and significant (0.133 with a standard error 0.032, and 0.037 with a standard error 0.018, respectively), however, the coefficient on $log(1+ExtSpill)$ is not significant (0.006 with a standard error of 0.021). The same pattern of results holds for a dynamic specification.

5.6. Robustness tests

Table 7 reports the robustness tests for the Tobin's Q estimation, using column 5 in table 4 as a benchmark.

The first robustness test relates to the concern that *IntSpill* and *ExtSpill* capture the patenting activity of the firm (instead of technological internalization), such that firms that have more patents will have higher *IntSpill* and lower *ExtSpill* (as the firm has more patents, the probability that it will randomly indirectly cite its previous patents is higher)[46]. The same pattern of results regarding *IntSpill* and *ExtSpill* can arise under this interpretation, if patents have a positive effect on private returns. In column 1, the citations-weighted patents stock is added. In case *IntSpill* and *ExtSpill* simply capture the patenting activity of the firm, they should be uninformative in this specification. I find the same pattern of results regarding *IntSpill* and *ExtSpill*, for the linear terms and the interacted terms. The coefficient on the interaction term of the citations-weighted patents stock is positive and significant, which is consistent with the higher appropriability interpretation.

The second robustness test relates to the size of the firm in the product market, aiming at mitigating the concern that larger firms are better at performing sequential innovation and capturing higher private returns. The same reasoning as for patenting activity is pursued, accordingly, *IntSpill* and *ExtSpill* should not be informative in the presence of product market size variables. As the linear terms of the size of the firm in the product market are already included, the R&D stock is interacted with Market Share. The same pattern of results regarding *IntSpill* and *ExtSpill* still remains.

Next, I test the robustness of the findings to including the classical measure of the external R&D stock of the firm (originally introduced by Jaffe, 1986), labeled as R&D Pool (appendix A.2 provides detail on its construction), as reported in column 3. Omitting R&D Pool may cause *IntSpill* to be upward biased and *ExtSpill* to be downward biased, from two main reasons. First, a higher R&D Pool can imply stronger competition in research.

[46] The correlation between *IntSpill* and citations-weighted patents stock is 0.173.

Thus, R&D Pool should have a negative effect on Tobin's Q. Under this interpretation, a higher R&D Pool should make it harder for the inventing firm to perform sequential innovation. Thus, technological internalization is upward biased. Second, a higher R&D Pool can imply stronger knowledge externalities, thus, the inventing firm can learn from the R&D that surrounds it. In this case, we should expect R&D Pool to positively affect Tobin's Q. Also, sequential innovation should be easier in case knowledge externalities intensify. Thus, technological internalization is overestimated in this case as well.

There is a positive and significant effect of R&D Pool on Tobin's Q via the interaction term with the R&D stock, and a negative, but not significant, linear effect. These results confirm the countervailing effects of R&D Pool on Tobin's Q: a positive learning effect through the R&D stock, and a negative linear competition effect[47]. However, technological internalization does not appear to be biased when R&D Pool is excluded, as there is no important change in $IntSpill$ and $ExtSpill$.

In column 4, citations-weighted patents stock, market share and R&D Pool (interacted and linear) are all included together. The same pattern of results regarding $IntSpill$ and $ExtSpill$ remains[48][49].

[Table 7 about here]

5.7. Estimation results for three specific industries

Since the above findings are based on a pooling estimation across industries, a big concern is that the diffusion variables capture variation in private returns across industries (instead of between-firms technological internalization). In order to mitigate this concern, I have controlled for industry effects, by including a complete set of two-digit industry dummies as

[47] Jaffe (1986) finds a similar negative linear effect of R&D Pool on market value, which he interprets as a negative competition effect in the technology space.

[48] I have also experimented with including self-citations (linearly and interacted with the R&D stock over assets), under the conjecture that self-citations represent an ability of the firm to conduct sequential innovation. In all specifications, the same pattern of results remains. With respect to self-citations, only the linear term is positive and significant.

[49] A linear version of equation (5.5) was also estimated with a complete set of four-digit SIC dummies. Only the coefficients on the interaction terms of $IntSpill$ and $ExtSpill$ remain significant (0.071 with a standard error of 0.021 and -0.008 with a standard error of 0.002, respectively).

a default in all specifications (in addition to main technology sector shares). Nevertheless, even if the cross-industry variation could be captured by two-digit industry dummies, they are included only linearly and not interacted with the R&D stock over assets (i.e., the linear industry effects are only an approximation of the cross-industry variation in private returns and technological internalization).

In this section, I estimate the effect of *IntSpill* and *ExtSpill* on Tobin's Q and R&D expenditures in three small panels, which differ in the importance they place on sequential innovation (appendix A.2 lists the industries that are included in each panel). Panel A includes all firms (from the sample of inventing firms) that operate in the "Semiconductors" industry. Sequential innovation plays a central role in this highly complex industry and I expect the diffusion variables to matter the most (this panel includes 12 firms that are active on average for 20 years)[50]. Panel B includes firms in the "Computers and Communications" industry (this sample includes 25 firms that are active on average for 20 years)[51]. Finally, panel C includes the firms in the "Drugs and Medicals" industry (this sample includes 19 firms that are active on average for 19 years). As sequential innovation plays a minor role in this industry, the diffusion variables should not matter much in this panel.

Table 8 summarizes the results, where the upper section reports the Tobin's Q estimation and the lower section reports the R&D estimation. Regarding Tobin's Q, the strongest results are in the "Semiconductors" panel. In this panel, the effect of the interaction term of *IntSpill* is positive and significant, where the effect of the interaction term of *ExtSpill* is negative and significant. When moving to the "Computers and Communications" panel, a similar pattern of results is observed. The interaction term of *IntSpill* is positive and significant, where the interaction term of *ExtSpill* is negative and significant (it is significant only in the specification with pre-sample means). Finally, in the "Drugs and Medicals" panel there is no significant effect of either *IntSpill* or *ExtSpill*.

[50]See Hall and Ziedonis (2001) for an analysis of the innovation and patenting activity in the "Semiconductors" industry.

[51]Note that panel A is a subset of panel B.

In the R&D specifications (static and dynamic) a similar pattern of results holds. Technological internalization matters for R&D expenditures in the "Semiconductors" and "Computers and Communications" panels, but does not matter in the "Drugs and Medicals" panel.

Overall, technological internalization matters the most in the industries where sequential innovation plays a central role. In these industries we should expect the theoretical predications to bind.

[**Table 8 about here**]

6. Summary and conclusions

This paper shows that firms are able to internalize dynamically some of their knowledge that spills to other firms. I exploit the firm-level variation of this internalization, through estimating the market valuation of the R&D stock of the firm. There is strong evidence suggesting that private returns rise with Internalized spillovers and fall with Externalized spillovers. Evaluated at the mean, a one standard deviation increase in $IntSpill$ raises the market valuation of an additional dollar spent on R&D by 30 percent, whereas a one standard deviation increase in $ExtSpill$ lowers the market valuation of an additional dollar spent on R&D by 10 percent.

In addition to quantifying the effect of technological internalization on private returns, the findings from the market value estimation also suggest that firms themselves are aware of their technological internalization and take it into consideration when making R&D decisions. To test this, a R&D equation is estimated, which shows that firms that create more Internalized and less Externalized spillovers, on average, invest substantially more in R&D (e.g., evaluated at the mean, a one standard deviation increase in $IntSpill$ raises R&D expenditures by 33 percent).

Finding that the R&D decision of the firm is affected by the pattern of diffusion its inventions follow has important implications for the endogenous growth literature. Spillovers encourage the innovation activity of the receivers of knowledge, however, their effect on the incentive to create knowledge at the first place depends on whether they are Internalized or Externalized. Suppose there are two economies that are identical in all dimensions, but in the first economy technological internalization is stronger. Based on the findings of this paper, innovation and, therefore, growth should be higher in the first economy

Moreover, the firm-level variation in technological internalization this paper has shown to exist may be linked to strategic behavior of firms optimizing the diffusion of their knowledge. To the extent technological internalization is subject to the behavior of firms has an important consequence for the way we model and think about knowledge spillovers.

7. References:

Aghion, P. and Howitt, P. (1992), "A Model of Growth through Creative Destruction", *Econometrica* Vol. 60, No. 2, pp. 323 351

Belenzon, S. (2005), "Knowledge Flows and Sequential Innovation: Implications for Technology Diffusion, Incentive for R&D and Firm Performance", *PhD Dissertation, University of London*

Bessen, K., Maskin, E. (2002), "Sequential Innovation, Patents and Imitation", *Department of Economics Working Paper 00-01, MIT*

Bloom, N., Schankerman, M. and Van Reenen, J. (2005), "Identifying technology spillovers and product market rivalry", *CEPR Discussion Papers 4912, C.E.P.R.*

Blundell, R., Griffith R. and Van Reenen, J. (1999), "Market share, Market value and Innovation in a Panel of British Manufacturing Firms", *Review of Economic studies*, 66, pp. 529-554

Caballero, R. J. and Jaffe, A. B. (1993), "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth", *NBER Working Papers 4370, National Bureau of Economic Research*

Chang, H. F. (1995), "Patent Scope, Antitrust Policy, and Cumulative Innovation", *RAND Journal of Economics*, Vol 26, pp. 34-57

Cockburn I., Griliches Z. (1988), "Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents", *American Economic Review*, Proceedings Issue Vol. 78, no. 2, pp. 419-423

Cornelli, F and Schankerman, M. (1999), "Patent Renewals and R&D Incentives", *RAND Journal of Economics*, pp. 17-34

Gilbert, R., Shapiro, C. (1990), "Optimal patent length and breadth", *RAND Journal of*

Economics, 21, pp. 106-112

Green, J., Scotchmer, S. (1995), "On the Division of Profit between Sequential Innovators", *The Rand Journal of Economics* 26, pp. 20-33

Griliches, Z. (1981), "Market Value, R&D and Patents", *Economics Letters*, 7, pp. 183-187

Griliches, Z. (1992), "The search for R&D spillovers", *Scandinavian Journal of Economics* 94, supplement, pp. S29-S47

Grossman, G. M. and Helpman, E. (1991), Innovation and Growth in the Global Economy, *MIT Press*, Cambridge, Mass

Hall, B. H., Jaffe A.B. , Trajtenberg M. (2001), "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools", *NBER Working Paper 8498*

Hall, B. H., Jaffe A.B. and Trajtenberg M. (2005), "Market Value and Patent Citations", *RAND Journal of Economics*, Vol. 36(1), pp. 16-38

Hall, B. H., Ziedonis Ham, Z. (2001), "The Determinants of Patenting in the U.S. Semiconductor Industry, 1980-1994," *Rand Journal of Economics,* 32, pp. 101-128

Jaffe, A. (1986), "Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits and market value", *American Economic Review,* 76, pp. 984-1001

Jaffe, A. and Trajtenberg, M. (1999), "International knowledge flows: Evidence from patent citations", *Economics of Innovation and New Technology*, Vol. 8, pp. 105-136

Jaffe, A. and Trajtenberg, M (2002), "Patents, Citations and Innovations: A Window on the Knowledge Economy", *Cambridge, Mass, MIT Press*, August 2002

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations", *Quarterly Journal of Economics* 108 (3), pp. 577-598

Jaffe, A., Fogarty, S., Trajtenberg, M. (2000), "Knowledge spillovers and Patent Citations:

Evidence from A Survey of Inventors", *American Economic Review*, pp. 215-218

Klemperer, P (1990), "How Broad Should the Scope of Patent Protection Be?", *RAND Journal of Economics*

Lanjouw, J. (1998), "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value", *The Review of Economics Studies*, Vol. 65, pp. 671-710

Lanjouw, J., Schankerman, M. (2004), "Patent quality and Research Productivity: Measuring Innovation with Multiple Indicators", *Economic Journal,* 114 (April), pp. 441-465

Pakes, A (1986), "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks", *Econometrica*, Vol. 54, no. 4,July, pp. 755-784

Romer, P. M. (1986), "Increasing Returns and Long-run Growth", *Journal of Political Economy*, Vol. 94(5), pp. 1002-37

Schankerman M. and Pakes A. (1986), "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period", *Economic Journal*, Vol.96, pp. 1052-1076

Schumpeter, J. A. (1942), "Capitalism, Socialism and Democracy", *New York: Harper and Brothers*

Scotchmer, S. (1991), "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law", Journal of Economic Perspectives 5(1), pp. 29-41

Scotchmer, S. (1996), "Protecting Early Innovators: Should Second-Generation Products be Patentable?", *The Rand Journal of Economics* 27, Summer, pp. 322-331

Scotchmer, S. (1999), "On the Optimality of the Patent Renewal System", *Rand Journal of Economics*, 30, pp. 181-196

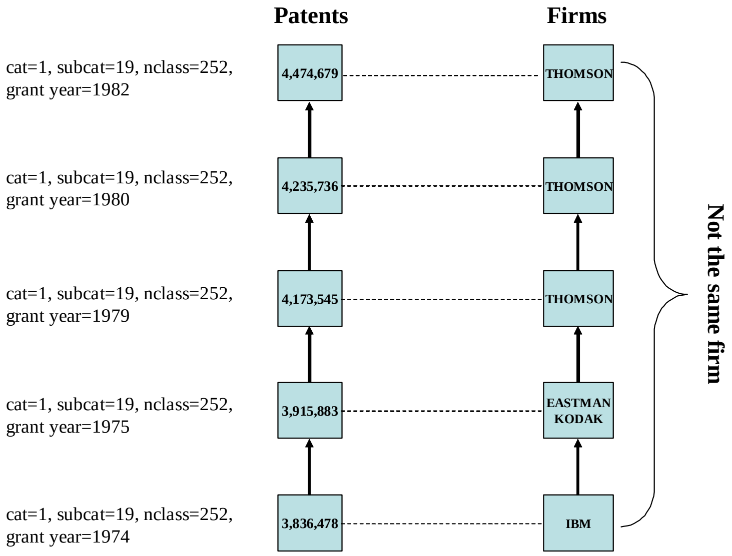
Segerstrom, P. S., Anant, C. A. and Dinopoulos, E. (1990), "A Schumpeterian Model of the Product Life Cycle", *American Economic Review*, 80, pp. 1077-1091

Trajtenberg, M. (1990), "A Penny for Your Quotes: Patent Citations and the Value of

Innovations", *The Rand Journal of Economics*, 21(1), pp. 172-187

Trajtenberg, M. (1990), "Economic Analysis of Product Innovation: The Case of CT Scanners", *Harvard University Press*,Cambridge, MassS63-S84

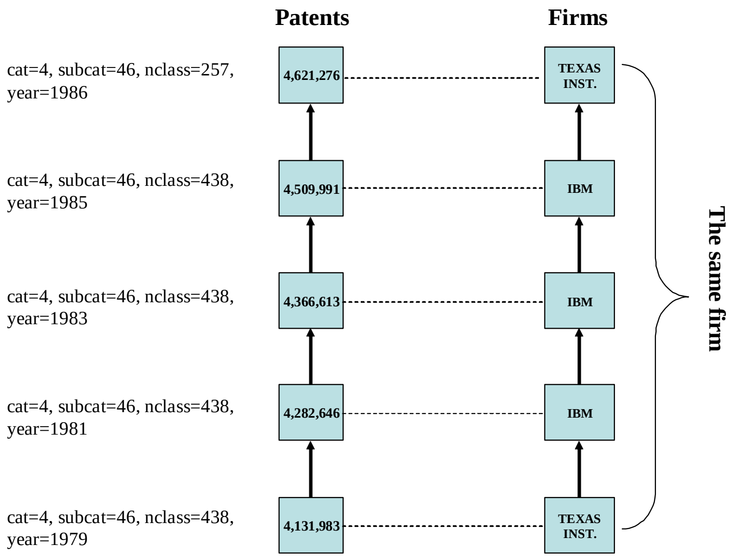
A. Appendices



cat refers to one-digit technology sectors, *subcat* refers to two-digits technology sectors and *nclass* refers to three-digits technology sectors.

Figure A1: An example for an Externalized line of research

Figure A1: *This figure shows a unique line of research originated in invention 3,836,478, which is owned by IBM (the inventing firm). Since knowledge did not return to IBM in the period 1974-1989, this line of research is Externalized.*



cat refers to one-digit technology sectors, *subcat* refers to two-digits technology sectors and *nclass* refers to three-digits technology sectors.

Figure A2: An example for an Internalized line of research

Figure A2: *This figure shows a unique line of research originated in invention 4,131,983, which is owned by Texas Instruments (the inventing firm). Since knowledge returned to Texas Instruments in the period 1979-1994, this line of research is Internalized.*

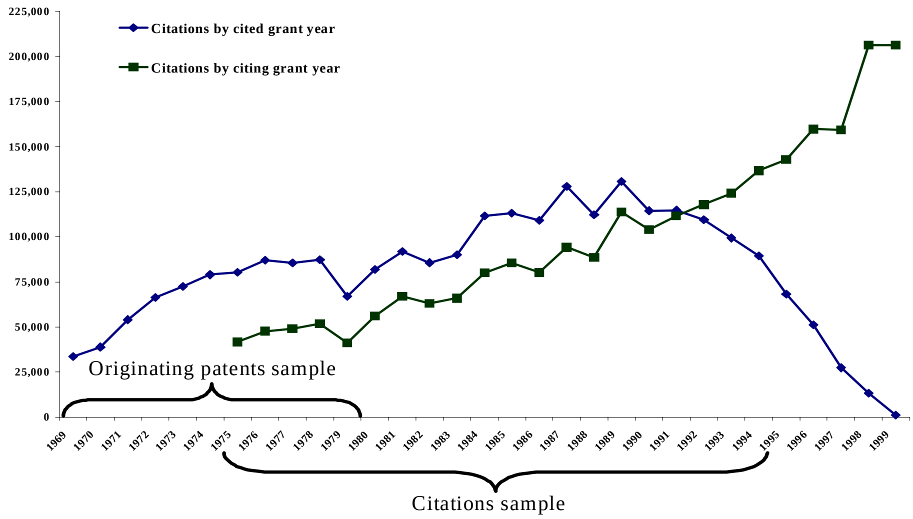


Figure A3: Citations sample

Figure A3: *This figure presents the number of citations made and received by patents in our sample. The upward sloping graph shows the number of citations made each year, where the U shaped curve shows the number of citations received each year.*

A.1. Generalizing proposition 2.1

This section shows that proposition (2.1) holds for every θ, i.e., R&D rises continuously with θ. The modified dynamic returns as a function of θ (following equation (2.6)) are:

$$W_i(\theta) = \frac{(1-q)\,p(x)v - x}{(1-p(x))\,(1-q)} \left(1 - \left(\frac{q}{1-(1-q)\,p(x)}\right)^{\theta+1}\right) \tag{A.1}$$

Since $W_{xx} < 0$ (the second derivative with respect to x), comparative statics imply:

$$\frac{dx}{d\theta} = sign(W_{x\theta}) \tag{A.2}$$

Differentiating equation $(A.1)$ with respect to x yields:

$$\begin{aligned} W_x &= \frac{(vp'(x)\,(1-q)-1)}{(1-q)\,(1-p(x))} \left(1 - \left(\frac{q}{1-p(x)\,(1-q)}\right)^{\theta+1}\right) \\ &\quad + \frac{p'(x)}{(1-p(x))} W(\theta) - p'(x)\,\frac{\theta+1}{1-p(x)} W(0) \left(\frac{q}{1-p(x)\,(1-q)}\right)^{\theta+1} \end{aligned} \tag{A.3}$$

And differentiating equation $(A.2)$ with respect to θ yields:

$$\begin{aligned} W_{x\theta} &= p'(x)\,\frac{W'(\theta)}{1-p(x)} - p'(x)\,\frac{W(0)}{1-p(x)} \left(\frac{q}{1-p(x)\,(1-q)}\right)^{\theta+1} \\ &\quad - p'(x)\,W(0)\,\frac{\theta+1}{1-p(x)} \left(\ln\frac{q}{1-p(x)\,(1-q)}\right) \left(\frac{q}{1-p(x)\,(1-q)}\right)^{\theta+1} \\ &\quad - \frac{(vp'(x)\,(1-q)-1)}{(1-q)\,(1-p(x))} \left(\ln\frac{q}{1-p(x)\,(1-q)}\right) \left(\frac{q}{1-p(x)\,(1-q)}\right)^{\theta+1} \end{aligned} \tag{A.4}$$

Since $\frac{q}{1-p(x)(1-q)} < 1$, for $W_{x\theta}$ to be positive it is enough to show that the following condition holds:

$$\left(\ln\frac{q}{1-p(x)\,(1-q)}\right) > \frac{1}{\theta+1} \tag{A.5}$$

For $\theta = 0$ condition $(A.5)$ holds as $q > 0$. Since the right-hand-side of condition $(A.5)$ decreases with θ, it must hold for every θ.

A.2. Data

The sample combines data mainly from two datasets:

The NBER USPTO patents database includes detailed patenting and citations information for around 2,600 US firms (as described in Hall, Jaffe and Trajtenberg (2001)) and a list of all the citations made in the period 1975-1999.

The Compustat North-America dataset provides full accounts data for over 25,000 US firms from 1980 to 2001. This provides information on the key accounting information of R&D, fixed assets, employment, sales, etc.

I started by matching the Compustat accounting data to the USPTO data, and kept firms with 1 or more patents in the period 1969-1980 that received at least one citation from the 2,600 firms in the NBER USPTO data set between 1975 and 1995. This leaves a sample of 512 firms. The accounting dataset has been 'cleaned' to remove accounting years with extremely large jumps (>+200% or <-66%) in sales, employment or capital signaling merger and acquisition activity, leaving 476 firms and a total of 9,454 observations.

A.2.1. The sample of originating patents

The set of originating patents (the set of inventions whose diffusion pattern is constructed) includes all cited patents that were granted between 1969 and 1980 and are held by the Compustat firms for which accounting data between 1980 and 2001 are available. The citations these patents receive must come from patents held by the 2,600 US Compustat firms between 1975 and 1995 (the set of citing patents includes 599,884 patents, which are about 30 percent of all citing patents and 50 percent of the US citing patents in the USPTO). The set of originating patents includes 104,694 patents.

Using about 1.7 million citations as technological links (where 599,884 patents cite 573,373 patents in the sample), 13,107,634 lines of research (singleton sequences of citations) are extracted, which are originated in 97,921 inventions[52]. 999,718 lines of research are classified as Internalized (7.6 percent of the total lines of research) and are originated in 29,964 patents (about 30 percent of the originating patents), while the remainder 12,107,916 lines of research are classified as Externalized and are originated in 97,212 patents[53].

A.2.2. Constructing the accounting variables

The book value of capital is the net stock of property, plant and equipment (Compustat Mnemonic PPENT); Employment is the number of employees (EMP). R&D (XRD) is used

[52] 6,773 patents that appear in our initial set of originating patents do not originate Internalized or Externalized lines of research. These patents originate lines of research in which all the follow-up developments of the originating invention is done within the boundaries of the inventing firm.

[53] The remaining 709 originating patents inspire only Internalized lines research (thus, all the subsequent generations of developments are done by the inventing firm).

to create R&D capital stocks calculated using a perpetual inventory method with a 15% depreciation rate (Hall et al, 2005). The citations-weighted patent stock was constructed by normalizing the number of patents the firm owns according to the number of citations it receives and the average number of citations to all patents in the same year. Given this normalized patents count the stock is constructed using the perpetual inventory method. The citations stock (used as a pre-estimation control) was constructed equivalently to the R&D stock. For Tobin's Q, firm value is the sum of the values of common stock, preferred stock, total debt net of current assets (Mnemonics MKVAF, PSTK, DT and ACT). Book value of capital includes net plant, property and equipment, inventories, investments in unconsolidated subsidiaries and intangibles other than R&D (Mnemonics PPENT, INVT, IVAEQ, IVAO and INTAN). Tobin's Q was set to 0.1 for values below 0.1 and at 20 for values above 20. See also Lanjouw and Schankerman (2004).

$R\&D\ Pool_{it}$ is constructed as:

$$R\&D\ Pool_{it} = \Sigma_{j,j\neq i} TEC_{ij}(R\&D\ Stock_{jt}) \tag{A.6}$$

Where, the index j represents firms that operate in overlapping technology sectors to firm i and TEC_{ij} is the classical measure of the level of orthogonality in research between firms i and j, originally developed by Jaffe (1986), and is defined as:

$$TEC_{ij} = \frac{(T_i T_j^{/})}{(T_i T_i^{/})^{\frac{1}{2}}(T_j T_j^{/})^{\frac{1}{2}}} \tag{A.7}$$

Where, T is a vector that its elements are the firm's share of patents in the three-digit technology sectors. The technology space information is provided by the allocation of all patents by the USPTO into 426 different technology classes. I use the average share of patents per firm in each technology class over the period 1970 to 2001 to create the following vector for each firm: $T_i = (T_{i,1}, T_{i,2}, ...T_{i,426})$, where $T_{i,m}$ is the share of patents of firm i in technology class m.

$Industry\ Sales_{it}$ is defined as the aggregate sales of other firms facing firm i (denoted by the index j), which operate in overlapping product markets, as following:

$$Industry\ Sales_{it} = \Sigma_{j,j\neq i} SIC_{ij}(Sales_{jt}) \tag{A.8}$$

Where, SIC_{ij} is defined following Bloom, Schankerman and Van Reenen (2005), as:

$$SIC_{ij} = \frac{(S_i S_j^{/})}{(S_i S_i^{/})^{\frac{1}{2}}(S_j S_j^{/})^{\frac{1}{2}}} \tag{A.9}$$

Where, S is a vector that its elements are the share of the firm's sales in the lines of business at the four-digit industry SIC codes. I use average share of sales per SIC code within each firm over the period as our measure of activity by product market,

$S_i = (S_{i,1}, S_{i,2}, ...S_{i,597})$, where $S_{i,m}$ is the share of sales of firm i in the four-digits SIC code m. The degree of orthogonality between every pair of firms is then computed (where higher orthogonality implies higher product market proximity). The normalization by the vector size aims to control for product diversity.

Industry price deflators were taken from Bartelsman, Becker and Gray, 2000, until 1996 and from the BEA 4-digit NAICS Shipment Price Deflators afterwards. Finally, Market Share is simply computed as the ratio between Sales and Industry Sales.

Finally, the industries considered in table 8 are as following: Semiconductors in panel (A) (which is a subset of panel (B)) covers only SIC 3674 (Semiconductors and Related Devices). Computers and Communications in panel (B) covers SIC 3571 (Electronic Computers), 3572 (Computer Storage Devices), 3661 (Telephone and Telegraph Apparatus), 3663 (Radio and Television Broadcasting and Communications Equipment), 3669 (Communications Equipment, Not Elsewhere Classified), 3674 (Semiconductors and Related Devices), 5065 (Electronic Parts and Equipment, Not Elsewhere Classified) and 5731 (Radio, Television, and Consumer Electronics Stores). Drugs and Medicals in panel (C) covers SIC 2834 (Pharmaceutical Preparations), 2835 (In Vitro and In Vivo Diagnostic Substances), 2844 (Perfumes, Cosmetics, and Other Toilet Preparations), 2851 (Paints, Varnishes, Lacquers, Enamels, and Allied Products), 3841 (Surgical and Medical Instruments and Apparatus), 3842 (Orthopedic, Prosthetic, and Surgical Appliances and Supplies), 3845 (Electromedical and Electrotherapeutic Apparatus) and 3851 (Ophthalmic Goods).

A.3. The Algorithm

This paper develops an algorithm that will generate a "family tree" for every originating patent in our sample. Since the computational task is highly complex and demanding, the efficiency of the algorithm plays a major role in making the task feasible. This section discusses the main steps of the algorithm. For the interested reader, a more detailed description is available upon request.

A.3.1. Source File

The source file contains the raw data, taken from the USPTO NBER Patents and Citations database. This file includes 1,760,143 rows, where each row corresponds to one patent citation, and 7 columns, which are the cited patent number, the citing patent number, the firm owning the cited patent, the firm owning the citing patent, the grant year of the cited patent, the grant year of the citing patent and an indicator to whether the cited patent is an originating patent.

The source file is sorted by the citing patent number. Thus, the first row is the earliest citation made in the sample, the second row is the second earliest citation etc. This sort allows saving valuable running time due to the fact that a citing patent cannot be cited before it cites. This sort is crucial for the running time of the algorithm.

A.3.2. Data Structure

In order to create an efficient algorithm that will produce the desired output in a reasonable time considering the amount of data, we use a combination of a Tree procedure and a Hash table. The Tree algorithm is a dynamic procedure that creates a 'tree' of patents without any restrictions on the number of both direct and indirect offspring patents. Each node in the 'tree' contains two types of information: information extracted from the source file, such as citing patent number and citing firm, and information that the algorithm generates, such as the location of the offspring patent in the 'tree'. Note that the 'tree' is not balanced (its branches are not of equal length), thus it does not benefit from the advantages of balanced 'trees', whose maximum length is already known. From this reason a Hash table is used, which allows us to efficiently store the information on the offspring patents in the diffusion 'tree' and save valuable search time.

The Hash table contains information on all the patents in the source file, both citing and cited defined as items. Each item contains the following fields: the depth in the 'tree' (the generation of citation), the place in the 'tree' (how it is linked to the originating patent) and an indictor to whether the patent is an originating patent. The place of the patent in the 'tree' is stored as a vector of numbers, as explained below.

A.3.3. Running process

Each row in the source file indicates a 'father-child' relationship in the 'tree'. The searching and updating procedure involves scanning the source file for every originating patent and updating the Hash table for each row according to the location of the citing patent in the diffusion 'tree' (in case a patent does not take part in the diffusion 'tree' of a given originating patent, its line is not updated).

The best way to explain the procedure of the algorithm is by a simple example. The following list of citing and cited patents is a sample taken from the source file:

Citing Patent	**Cited Patent**
3988245	3852388
3988250	3852388
4032309	3852388
4119408	4032309
4174374	4119408
4564373	4174374
4617029	4174374
4629563	3988245
4629570	3988255
4666607	3988245
4737166	4174374

Given this list, the algorithm will begin with the first row in the file, which says that patent number 3988245 cites patent number 3852388. As the algorithm starts to construct a new diffusion 'tree', it first checks whether the cited patent in the first row is part of the set of the originating patents. If it is not part of this set, the algorithm skips this row and jumps to the next one. If it does belong to the set of originating patents, the algorithm starts the construction of the diffusion 'tree' for this patent by updating the Hash table for this row and for the next rows in the source file. We will show now how the updating procedure takes place.

The entries in the Hash table at the end on the running and updating procedure is as following (at the beginning of the procedure, the items in the Hash table are initialized to -1):

Patent number	Originating	Place	Depth
...			
3852388	$originating = 1$	$PlaceInTree = 1$	$Depth = 1$
3988245	$originating = 0$	$PlaceInTree = 11$	$Depth = 2$
3988250	$originating = 0$	$PlaceInTree = 12$	$Depth = 2$
4032309	$originating = 0$	$PlaceInTree = 13$	$Depth = 2$
4119408	$originating = 0$	$PlaceInTree = 131$	$Depth = 3$
4174374	$originating = 0$	$PlaceInTree = 1311$	$Depth = 4$
4564373	$originating = 0$	$PlaceInTree = 13111$	$Depth = 5$
4617029	$originating = 0$	$PlaceInTree = 13112$	$Depth = 5$
4629563	$originating = 0$	$PlaceInTree = -1$	$Depth = -1$
3988255	$originating = 0$	$PlaceInTree = -1$	$Depth = -1$
4629570	$originating = 0$	$PlaceInTree = -1$	$Depth = -1$
4666607	$originating = 0$	$PlaceInTree = -1$	$Depth = -1$
4737166	$originating = 0$	$PlaceInTree = 13113$	$Depth = 5$
...			

Once the algorithm finishes scanning the source file, another function is called in to print all the branches of the 'tree' into a file. These branches are unique sequences of patent citations, which we interpret as lines of research. The printed lines of research are than given in a text format ready to be analyzed in any statistical package. Determining whether a line of research is Internalized or Externalized is a straightforward task, as we only need to compare the first firm in the sequence of citations to the last firm. If these are identical (and there is at least one external invention along the line of research, such that spillovers are created), the line of research is Internalized. Other wise, it is classified as Externalized.

The next step is to clean the memory and initialize the Hash table before proceeding to the next originating patent, and repeating the same algorithm.

A.4. Theoretical model - developing equation (2.6)

This section shows how the expression of the dynamic returns in equation (2.6) is derived.

The model does not include time, only generation of developments of the originating knowledge k. Suppose the model starts at the point in time where knowledge k becomes available for sequential innovation by other firms (and by the inventing firm i). All computations of the expected number of wins relate to the point of view of this starting period.

The probability that firm i wins in generation g, as calculated at the initial period, is:

$$P(g) = \sum_{s=0}^{g-1} \binom{g-1}{s} [p(1-q)]^{g-s} q^s \tag{A.10}$$

It should be noted that the term q^s reflects the ability of the firm to build on external research along the line of research it originates. The probability that knowledge is created in a given development stage and firm i not winning in this stage is $q(1-p) + pq = q$ (since the firm does not win either if it fails to invent, or if it succeeds to invent, however, at least one other firm succeeds as well).

I aim at computing the expected dynamic returns to knowledge k, given the expected number of development stages won by firm i. For this purpose, the following equation for the expected number of development stages won by firm i has to be computed:

$$E(wins) = \sum_{g=0}^{\infty} P(g) = \sum_{g=0}^{\infty} \sum_{s=0}^{g-1} \binom{g-1}{s} [p(1-q)]^{g-s} q^s \tag{A.11}$$

Taking g to infinity (assuming the knowledge k has the potential of being developed an infinite number of times) and computing the expected number of inventions firm i makes along the line of research can be expressed as following:

$$\begin{array}{lllll}
p(1-q) & & & & \\
p^2(1-q)^2 & p(1-q)q & & & \\
p^3(1-q)^3 & 2p^2(1-q)^2q & p(1-q)q^2 & & \\
p^4(1-q)^4 & 3p^3(1-q)^3q & 3p^2(1-q)^2q^2 & p(1-q)q^3 & \\
p^5(1-q)^5 & 4p^4(1-q)^4q & 6p^3(1-q)^3q^2 & 4p^2(1-q)^2q^3 & p(1-q)q^4
\end{array}$$

Define $h \equiv (1-q)\,p$. The summation of equation (5.23) over g can be computed by first summing each column across its rows and then summing over columns. Also, define s as the number of times the firm failed to win a development stage and then summing over s equals 0 to infinity.

Summation of $s = 0$ (zero failures):

$$S^0 = h + h^2 + h^3 + h^4 + ... \tag{A.12}$$

$$S^0 = \frac{h}{1-h} \tag{A.13}$$

Summation of $s = 1$ (one failure):

$$S^1 = q\left(h + 2h^2 + 3h^3 + 4h^4 + ...\right) \tag{A.14}$$

Which can be written, as following:

$$\begin{array}{lllll} h & h^2 & h^3 & h^4 & ... \\ & h^2 & h^3 & h^4 & ... \\ & & h^3 & h^4 & ... \\ & & & h^4 & ... \end{array}$$

Using the same method, I can first sum across rows and then across columns. This yields:

$$S^1 = \frac{q}{1-h}\left[h + h^2 + h^3 ...\right] = \frac{q}{1-h} S^0 \tag{A.15}$$

For $s = 2$ (two failures) I get the following summation:

$$S^2 = q^2\left(h + 3h^2 + 6h^3 + 10h^4 + ...\right)$$

Which can be expressed in the following form:

$$\begin{array}{lllll} h & h^2 & h^3 & h^4 & ... \\ & h^2 & h^3 & h^4 & ... \\ & h^2 & h^3 & h^4 & ... \\ & & h^3 & h^4 & ... \\ & & h^3 & h^4 & ... \\ & & h^3 & h^4 & ... \\ & & & h^4 & ... \\ & & & h^4 & ... \\ & & & h^4 & ... \\ & & & h^4 & ... \end{array}$$

Using the same method described above, this summation becomes:

$$S^2 = \frac{q}{1-h} q\left(h + 2h^2 + 3h^3 + 4h^4 ...\right) = \frac{q}{(1-h)} S^1 \tag{A.16}$$

With $s = 3$ (three failures) the summation is:

$$S^3 = q^3\left(h + 4h^2 + 10h^3 + ...\right) \tag{A.17}$$

Which can be expressed, as following:

$$\begin{array}{llll} h & h^2 & h^3 & ... \\ & h^2 & h^3 & ... \\ & h^2 & h^3 & ... \\ & h^2 & h^3 & ... \\ & & h^3 & ... \\ & & h^3 & ... \\ & & h^3 & ... \\ & & h^3 & ... \\ & & h^3 & ... \\ & & h^3 & ... \end{array}$$

As before, this summation becomes:

$$S^3 = \frac{q}{1-h} q^2 \left(h + 3h^2 + 6h^3 ...\right) = \frac{q}{(1-h)} S^2 \tag{A.18}$$

Thus, the summation of columns is a geometric series with a multiplicative factor equals $\frac{q}{(1-h)}$ and the first argument in the series is $\frac{h}{1-h}$.

Thus, the dynamic returns as a function of the number of 'second chances' the firm gets, θ, are given as (thus, θ is the number of columns to sum:

$$W_i(\theta) = \frac{(1-q)\,pv - x}{(1-p)\,(1-q)} \left(1 - \left(\frac{q}{1-(1-q)\,p}\right)^{\theta+1}\right) \tag{A.19}$$

A.5. The technological contribution of an invention - an alternative interpretation

This section shows that the methodology of measuring technological contribution is a generalization of the accepted approach of measuring the quality of patents by counting the number of citations they receive, while also including indirect offspring patents.

To illustrate, refer back to figure 2. Under pattern 1 there are three offspring inventions, and, therefore, there are three citing (direct and indirect) patents, where each citing patent is cited only once. It should be noted that it is assumed that the last patent in the sequence is counted as if it receives one citation. Thus, $TC_A^1 = (1 \times 1) + (1 \times 1) + (1 \times 1) = 3$. With respect to diffusion pattern 2, there are three offspring inventions as well. However, patent B receives two direct citations and, therefore, it receives the weight of 2. This implies that $TC_A^1 = (1 \times 2) + (1 \times 1) + (1 \times 1) = 4$. These measures are identical to the lines of research approach.

A closer look at this methodology would show that the scheme is recursive. Assume patent C in figure 2 under diffusion pattern 1 receives another citation from patent E (thus, it is cited twice, by patent D and patent E). Using the lines of research approach, there are two lines of research: $A \to B \to C \to D$ and $A \to B \to C \to E$. Thus, $TC_A^1 = (1 \times 3) + (1 \times 3) = 6$. Under the alternative approach discussed above, TC_A^1 is computed as following: starting with patent C (we continue to assume that the edge patents, D and E in this case, are cited only once), it receives two citations of quality one each (the quality of patents D and E). Regarding patent B, it is cited only once (by patent C). However, since patent C is of quality 2, I treat the citation from patent C to patent B, as if patent B receives two citations. In this case, $TC_A^1 = (1 \times 2) + (1 \times 2) + (1 \times 1) + (1 \times 1) = 6$, which is the same as the technological contribution under the lines of research approach.

More formally, the alternative interpretation of the methodology is the following:

$$TC_i = \sum_{k \in K_i} OS_{ik} \times \widehat{Q}_k \tag{A.20}$$

And $\widehat{Q}_k$ is expressed as:

$$\widehat{Q}_k = \sum_{j \in J} OS_{kj} \times \widehat{Q}_j \tag{A.21}$$

Where, K_i is the set of patents that cite directly or indirectly patent i, OS_{ik} denotes the offspring invention $k \in K_i$, j is another patent in the set K_i, which directly cites invention k, J_k is the set of patents that directly cite invention k (i.e., $j \in J_k \subset K_i$), OS_{kj} denotes the offspring invention which directly cite patent k and $\widehat{Q}_j$ is the quality of invention j.

A.6. Linearizing equation (5.5)

A linear version of equation (5.5) is estimated to test the robustness of the findings to specifications where the term $log(1+\gamma\frac{K_{it}}{A_{it}})$ is approximated by a polynomial series expansion. The series of functions used for this approximation is denoted by $\gamma\Phi(\frac{K_{it}}{A_{it}})$, which is linear in γ. I experiment with a series expansion of degree one ($\Phi(\frac{K_{it}}{A_{it}}) = \frac{K_{it}}{A_{it}}$), two ($\Phi(\frac{K_{it}}{A_{it}}) = \sum_{j=1}^{2}\left(\frac{K_{it}}{A_{it}}\right)^{j}$), three ($\Phi(\frac{K_{it}}{A_{it}}) = \sum_{j=1}^{3}\left(\frac{K_{it}}{A_{it}}\right)^{j}$) and four ($\Phi(\frac{K_{it}}{A_{it}}) = \sum_{j=1}^{4}\left(\frac{K_{it}}{A_{it}}\right)^{j}$). Thus, equation (5.3) becomes:

$$\log\left(\frac{V_{it}}{A_{it}}\right) = \log\kappa_{it} + \gamma\Phi(\frac{K_{it}}{A_{it}}) \tag{A.22}$$

Where, γ and $\log\kappa_{it}$ are specified in equations (5.2) and (5.4) , respectively. Equation (5.5) is estimated by OLS, where the standard errors of the marginal effects are computed using the Delta method.

Table A3 reports the estimation results with a series expansion of degree one (thus, $log(1+\gamma\frac{K_{it}}{A_{it}}) \approx \gamma\frac{K_{it}}{A_{it}}$). Column 1 reports the estimation results of including R&D over assets linearly and interacted with *IntSpill* and *ExtSpill* (as in column 1 in table 4 for the nonlinear specification). The pattern of results is similar to the one observed in the nonlinear estimation. The coefficient on the interaction term of *IntSpill* is positive and significant and the coefficient on the interaction term of *ExtSpill* is negative and significant. Compared to the equivalent nonlinear specification (column 2 in table 4), the linear specification yields a lower coefficient on the interaction term of *IntSpill* (0.089 compared to 0.208) and a lower coefficient on the R&D over assets (0.229 compared to 0.280). The coefficient on the interaction term of *ExtSpill* is similar to the coefficient obtained from the nonlinear specification.

The elasticity of market value with respect to the R&D stock, evaluated at the mean, is 0.093, compared to 0.103 in the equivalent nonlinear specification. An additional one dollar spent on R&D raises market value by 0.44 dollar, compared to 0.49 dollar in the nonlinear specification. A one standard deviation increase in $IntSpill$ raises private returns by 17 percent (compared to 30 percent in the nonlinear specification), whereas a one standard deviation increase in $ExtSpill$ lowers private returns by 5 percent (compared to 10 percent in the nonlinear specification)[54].

In column 2, the set of pre-sample means is added. The coefficients on R&D stock over assets and the interaction terms of *IntSpill* and *ExtSpill* substantially drop, however, their signs do not change and they remain significant. In column 3, I add the linear terms

[54] The effect of $IntShare$ in the linear specification is identical to its effect in the nonlinear specification. A one standard deviation increase in $IntShare$ raises private returns to an extra dollar spent on R&D by 39 percent.

of *IntSpill* and *ExtSpill*. The coefficients on the linear terms of *IntSpill* and *ExtSpill* are similar to those obtained from the nonlinear specification, i.e., *IntSpill* is positive and significant and *ExtSpill* is negative and significant. The coefficient on the interaction term of *IntSpill* drops, however it remains significantly positive, while the coefficient on the interaction term of *ExtSpill* do not change much.

In columns 4 to 6, similar robustness tests are repeated as reported in table 5. Thus, adding linearly and interacted the citations-weighted patents stock and R&D Pool. The same pattern of results regarding *IntSpill* and *ExtSpill* (linear and interacted) remains.

[**Table A3 about here**]

In table A4, I experiment with higher degrees of polynomial approximation[55]. Columns 2 to 4 report the estimation results with a polynomial expansion of degree two, three and four, respectively. I find the same pattern of results regarding the interaction terms of *IntSpill* and *ExtSpill*. With regard to the coefficients size, as the degree of the polynomial expansion rises, the effects of the linear term of R&D stock over assets and its interactions with *IntSpill* and *ExtSpill* rise. For example, the elasticity of Tobin's Q with respect to R&D stock over assets in the fourth-degree polynomial approximation is 0.14, compared to 0.09 in the second-degree polynomial approximation. Nevertheless, although the size of the effect changes, the pattern of results is very robust to any form of linear approximation.

[**Table A4 about here**]

Finally, I test the robustness of the above findings to four-digit industry effects. Table A5 reports the estimation results of including a complete set of four-digit industry dummies in polynomial expansions of degrees one and two. The same pattern of results holds, where the coefficient on *IntSpill* is positive and significant and the coefficient on *ExtSpill* is negative and significant. Interestingly, the effects of the linear term of R&D stock over assets and its interactions with *IntSpill* and *ExtSpill* rise when exploiting only the variation within four-digit industry SIC codes.

[**Table A5 about here**]

[55] The marginal effects are computed by differentiating equation (5.5) with respect to each variable. Standard errors for the marginal effects are computed using the Delta method.

Table 1

Internalized and Externalized lines of research

	Number of lines of research[a]	Share of Internalized lines of research[b]			
	Total sample	Total sample	1969-1975	1976-1978	1979-1980
Pooled	46.8	7.6%	8.2%	7.6%	7.2%
Chemicals	28.8	6.2%	6.4%	6.3%	5.7%
Computers and Communications	30.2	7.6%	8.8%	7.1%	7.1%
Drugs and Medicals	16.8	15.0%	19.1%	16.8%	8.4%
Electrical and Electronics	78	7.4%	7.5%	7.1%	7.5%
Mechanicals	15.5	8.8%	9.1%	9.1%	7.9%

[a]Computed as the average number of lines of research per citations received by an originating patent for the entire period of the sample.

[b]Computed as the ratio between Internalized lines of research and the total number of lines of research.

Table 2

Descriptive statistics: accounting and patents variables

9,454 observations and 476 firms

Variable	Mnemonic	Mean	Median	Min	Max	Standard deviation
IntSpill		5.90	0.14	0	891	45
ExtSpill[1]		4.28	0.38	0.01	391.50	25.2
IntShare[2]		0.02	0.00	0	0.25	0.04
Tobin's Q	V/A	2	1.32	0.1	20	2.34
Market value, $m	V	4,689	592	0	485,566	16,782
R&D stock, $m	K	806	49	0	47343	3195
R&D stock / Assets	K/A	0.39	0.20	0	10	1
Capital stock, $m	A	3,090	392	2.13	199,303	9,736
Sales, $m		3,925	686	0	180,557	11,412
Patents stock		155	18	0.42	9,848	489
Patents stock weighted by citations		158	16	0.28	12,643	585

The statistics are computed over all the observations that were included in the estimation (1980-2001) and are given in thousands of 1996 USD.

[1]Divided by 100.

[2]For about 40 percent of firms IntSpill equals zero.

Table 3

Analysis of Variance - diffusion variables

	One-digit SIC	Two-digits SIC	Three-digits SIC	Four-digits SIC
IntSpill	1.51	0.30	0.23	0.24
% Between industries variation	3%	6%	8%	15%
% Within industries variation	97%	94%	6%	85%
ExtSpill	2.84*	0.54	0.36	0.61
% Between industries variation	5%	9%	12%	31%
% Within industries variation	95%	91%	88%	69%
IntShare	1.37	1.71*	1.47*	1.52*
% Between industries variation	3%	25%	36%	52%
% Within industries variation	97%	75%	64%	48%

Table entries are the F-statistics for the null hypothesis of equal mean across the different industry breakdowns. * denotes that the mean varies across industries at the 5 percent significance level.

Table 4

The effect of IntSpill and ExtSpill on Tobin's Q

Nonlinear Least Squares, dependent variable: log(Tobin's-Q)

	(1)	(2)	(3)	(4)	(5)
R&D stock/Assets	0.280*	0.145*	0.152*	0.167*	0.217*
	(0.079)	(0.021)	(0.028)	(0.029)	(0.040)
IntSpill x (R&D stock/Assets)	0.208*	0.096*	0.059*	0.059*	0.062*
	(0.095)	(0.029)	(0.010)	(0.009)	(0.012)
ExtSpill x (R&D stock/Assets)	-0.011*	-0.005*	-0.004*	-0.004*	-0.005*
	(0.003)	(0.002)	(0.001)	(0.001)	(0.002)
log(IntSpill)			0.031*	0.027*	0.028*
			(0.005)	(0.005)	(0.005)
log(ExtSpill)			-0.023*	-0.026*	-0.026*
			(0.004)	(0.004)	(0.004)
log(Sales)				0.031*	0.028*
				(0.003)	(0.003)
log(Industry Sales)				-0.024*	-0.029*
				(0.006)	(0.006)
Sales Growth					0.533*
					(0.017)
Pre-sample means[a]	No	Yes	Yes	Yes	Yes[b]
Observations	9,454	9,454	9,454	9,454	9,015
R^2	0.323	0.501	0.509	0.511	0.516

Standard errors (in brackets) are robust to arbitrary heteroskedacity and serial correlation (clustered at the firm level). * denotes a significant level of 5 percent.

All regressions include 78 two-digits industry dummies, 4 technology indicators, a complete set of year dummies, a dummy variable for R&D stock equals zero and a dummy variable for IntSpill equals zero.

[a]The set of pre-sample means: Market Share, Employees, Tobin's Q, Sales, Assets, R&D stock, Patents stock and Citations stock.

[b]The estimates for the pre-sample mean variables are as following: Employees 0.003 (0.002), Market Share -0.106* (0.018), log(Tobin's Q) 0.544* (0.011), Sales 0.013 (0.027), Assets -0.856 (0.592), Patents stock -0.009 (0.013), Citations stock 0.029* (0.007) and R&D stock -0.292 (0.335).

Table 5

Quantitative effects of IntSpill, ExtSpill and IntShare

	Interaction terms	Firm fixed-effects	Linear effects
	(1)	(2)	(3)
One standard deviation increase			
Internalized Flows	+30%	+30%	+16%
Externalized Spillovers	-10%	-10%	-6%
Internalized Share	+40%	+50%	+37%

Note: columns (1), (2) and (3) are based on the corresponding columns in tables 3 and A5. Thus, column 1 includes only R&D stock over assets and interactions with IntSpill, ExtSpill and IntShare, column 2 adds pre-sample means and column 3 adds the linear terms of IntSpill, ExtSpill and IntShare.

Table 6

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
R&D expenditures and firm fixed-effects (FE[1]) estimation										
9,454 observations (in the R&D regression), 476 firms (for the fixed-effects regressions)										
	log(R&D)		Fixed-effcets							
	Static	Dynamic	Static		dynamic		static		dynamic	
log(IntSpill)			0.276*	0.197*	0.182*	0.083*				
			(0.030)	(0.031)	(0.088)	(0.038)				
log(ExtSpill)			-0.048*	-0.032*	-0.032*	-0.019*				
			(0.014)	(0.011)	(0.011)	(0.009)				
Dummy for IntSpill=0			-2.619*	-2.083*	-0.901*	-0.736*				
			(0.170)	(0.171)	(0.079)	(0.074)				
IntShare							11.862*	8.075*	7.669*	5.755*
							(2.000)	(1.392)	(1.423)	(1.050)
mean(Sales)				0.006		0.022		0.033*		0.047*
				(0.011)		(0.013)		(0.001)		(0.012)
mean(Industry Sales)				0.053*		0.034*		0.041*		0.035*
				(0.008)		(0.006)		(0.008)		(0.006)
mean(CW Patents stock)				0.015		0.019		0.011*		0.054*
				(0.018)		(0.021)		(0.003)		(0.014)
mean(R&D stock)				-0.041		-0.008		0.002*		-0.015
				(0.045)		(0.054)		(0.001)		(0.009)
log(R&D$_{t-1}$)		0.568*								
		(0.006)								
log(Sales$_t$)	0.146*	0.262*								
	(0.013)	(0.007)								
log(Sales$_{t-1}$)	0.044*	-0.199*								
	(0.008)	(0.007)								
log(Industry Sales$_t$)	0.296*	0.125*								
	(0.019)	(0.014)								
log(Industry Sales$_{t-1}$)	-0.033	-0.007								
	(0.023)	(0.008)								
log(CW Patents stock)	0.286*	0.079*								
	(0.011)	(0.008)								
R^2	0.778	0.949	0.391	0.543	0.223	0.458	0.113	0.425	0.077	0.379

[1]The estimated firm fixed-effects are fitted from the R&D regression that is reported in column 1.
Robust standard errors are in brackets. * denotes a significance level of 5 percent.

The R&D equation includes a complete set of year dummies and a dummy for R&D equals zero.

Table 7

Robustness tests for the Tobin's Q estimation

Nonlinear Least Squares, dependent variable: log(Tobin's-Q)

	(1)	(2)	(3)	(4)
R&D stock/Assets	0.204*	0.222*	0.183*	0.187*
	(0.039)	(0.007)	(0.037)	(0.038)
IntSpill x (R&D stock/Assets)	0.048*	0.063*	0.062*	0.052*
	(0.013)	(0.012)	(0.012)	(0.014)
ExtSpill x (R&D stock/Assets)	-0.004*	-0.005*	-0.005*	-0.005*
	(0.001)	(0.002)	(0.002)	(0.002)
log(IntSpill)	0.025*	0.028*	0.024*	0.021*
	(0.005)	(0.005)	(0.005)	(0.005)
log(ExtSpill)	-0.027*	-0.026*	-0.026*	-0.027*
	(0.004)	(0.004)	(0.004)	(0.004)
CW Patents Stock x (R&D stock/Assets)	0.006*			0.004*
	(0.002)			(0.002)
log(CW Patents Stock)	0.010			0.016*
	(0.006)			(0.006)
Market Share x (R&D stock/Assets)		-0.035		-0.058
		(0.066)		(0.062)
Market Share		0.061		0.073
		(0.049)		(0.051)
R&D Pool x (R&D stock/Assets)			0.074	0.059
			(0.064)	(0.066)
log(R&D Pool)			-0.018*	-0.024*
			(0.009)	(0.009)
Pre-sample means[a]	Yes	Yes	Yes	Yes
Observations	9,015	9,015	9,015	9,015
R^2	0.516	0.516	0.516	0.516

Standard errors (in brackets) are robust to arbitrary heteroskedacity and serial correlation (clustered by firms). * denotes a significant level of 5 percent.

All regressions include 78 two-digits industry dummies, 4 technology indicators, a complete set of year dummies, a dummy variable for R&D stock equals zero, a dummy variable for IntSpill equals zero, sales and industry sales.

[a]The set of pre-sample means: Market Share, Employees, Tobin's Q, Sales, Assets, R&D stock, Patents stock and Citations stock.

Table 8

The effect of IntSpill and ExtSpill - Heterogeneous industries

Dependent variable: log(Tobin's Q)						
	(A) Semiconductors[1]		(B) Computers and Communications[2]		(C) Drugs and Medicals[3]	
R&D stock/Assets	0.442	0.022	0.319*	0.133*	0.133	0.117*
	(0.367)	(0.146)	(0.095)	(0.062)	(0.109)	(0.057)
IntSpill x (R&D stock/Assets)	0.057*	0.047*	0.089*	0.047*	0.237	0.340
	(0.023)	(0.012)	(0.015)	(0.012)	(0.542)	(0.229)
ExtSpill x (R&D stock/Assets)	-0.004	-0.008*	-0.018*	0.001	0.055	-0.435
	(0.007)	(0.004)	(0.005)	(0.004)	(0.922)	(0.368)
Pre-sample means	No	Yes	No	Yes	No	Yes
Observations	240	240	501	501	357	357
R^2	0.308	0.516	0.189	0.448	0.207	0.567

Standard errors are robust to arbitrary heteroskedacity and serial correlation. * denotes a significance level of 5 percent.

Dependent variable: Firm fixed-effects from the first stage R&D equation						
	(A) Semiconductors[1]		(B) Computer and Communications[2]		(C) Drugs and Medicals3	
	R&D Static	R&D Dynamics	R&D Static	R&D Dynamics	R&D Static	R&D Dynamics
IntSpill	0.279*	0.318*	0.283*	0.189*	0.031	-0.009
	(0.092)	(0.089)	(0.136)	(0.095)	(0.238)	(0.266)
ExtSpill	-0.029	-0.054*	-0.012	-0.008	0.221	-0.049
	(0.021)	(0.019)	(0.039)	(0.033)	(1.030)	(1.253)
Dummy for IntSpill=0	-2.098*	-2.323*	-1.748*	-1.316*	-2.126*	0.055
	(0.729)	(0.646)	(0.652)	(0.436)	(1.061)	(1.177)
Observations	12	12	25	25	19	19
R^2	0.629	0.722	0.484	0.402	0.422	0.432

Standard errors are robust to arbitrary heteroskedacity. * denotes a significance level of 5 percent.

The R&D estimation is based on column 2 in table 6 for the static specification and column 2 in table A4 for the dynamic specification (including the lag of R&D on the right-hand-side of the first-stage estimation).

Table A1

The diffusion variables and the main characteristics of the firm: OLS estimation

	IntSpill	ExtSpill	IntShare
log(mean Sales)	4.162	0.338	-0.002
	(3.585)	(0.209)	(0.003)
log(mean R&D Stock)	-0.127	0.064	0.000
	(1.138)	(0.066)	(0.001)
log(mean Employees)	-5.079	-0.613	0.001
	(3.994)	(0.233)	(0.003)
log(mean CW Patents Stock)	5.274	0.034	0.009
	(1.873)	(0.109)	(0.001)
log(mean Citations Stock)	1.075	0.259	-0.001
	(2.188)	(0.128)	(0.002)

The estimation sample includes the 476 firms that are in our final sample.
Standard errors (in brackets) are robust to arbitrary heteroskedacity and serial correlation. * denotes a significant level of 5 percent.

Table A2

The effect of InteShare on Tobin's Q

Nonlinear Least Squares, dependent variable: log(Tobin's-Q)

	(1)	(2)	(3)	(4)	(5)
R&D stock/Assets	0.330*	0.120*	0.135*	0.141*	0.217*
	(0.101)	(0.064)	(0.026)	(0.026)	(0.040)
IntShare x (R&D stock/Assets)	5.624*	2.341*	1.702*	1.379*	1.311*
	(2.295)	(0.507)	(0.533)	(0.498)	(0.586)
log(IntShare)			0.016*	0.020*	0.025*
			(0.004)	(0.006)	(0.007)
log(Sales)				0.035*	0.033*
				(0.004)	(0.004)
log(Industry Sales)				-0.005*	-0.011*
				(0.006)	(0.006)
Sales Growth					0.538*
					(0.018)
Pre-sample means[a]	No	Yes	Yes	Yes	Yes
Observations	9,454	9,454	9,454	9,454	9,015
R^2	0.294	0.496	0.496	0.499	0.504

Standard errors (in brackets) are robust to arbitrary heteroskedacity and serial correlation (clustered at the firm level). * denotes a significant level of 5 percent.

All regressions include 78 two-digits industry dummies, 4 technology indicators, a complete set of year dummies, a dummy variable for R&D stock equals zero and a dummy variable for IntSpill equals zero.

[a]The set of pre-sample means: Market Share, Employees, Tobin's Q, Sales, Assets, R&D stock, Patents stock and Citations stock.

Table A3

The effect of IntSpill and ExtSpill on Tobin's Q

Linear estimation, dependent variable: log(Tobin's-Q)

	(1)	(2)	(3)	(4)	(5)	(6)
R&D stock/Assets	0.229* (0.022)	0.137* (0.025)	0.141* (0.025)	0.185* (0.028)	0.187* (0.038)	0.174* (0.037)
IntSpill x (R&D stock/Assets)	0.089* (0.022)	0.044* (0.015)	0.026* (0.012)	0.030* (0.015)	0.029* (0.015)	0.028* (0.014)
ExtSpill x (R&D stock/Assets)	-0.008* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)
log(IntSpill)			0.032* (0.008)	0.022* (0.008)	0.031* (0.008)	0.020* (0.009)
log(ExtSpill)			-0.018* (0.007)	-0.026* (0.004)	-0.019* (0.007)	-0.021* (0.007)
CW Patents Stock x (R&D stock/Assets)				0.003 (0.020)		0.003 (0.020)
log(CW Patents Stock)				0.033* (0.009)		0.016* (0.006)
R&D Pool x (R&D stock/Assets)					0.002 (0.032)	0.002 (0.033)
log(R&D Pool)					-0.018 (0.014)	-0.043* (0.015)
Sales Growth				0.556* (0.047)	0.557* (0.048)	0.548* (0.048)
Observations	9,454	9,454	9,454	9,015	9,015	9,015
Pre-sample means[a]	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors (in brackets) are robust to arbitrary heteroskedacity and serial correlation (clustered by firms). * denotes a significant level of 5 percent.

All regressions include 78 two-digits industry dummies, 4 technology indicators, a complete set of year dummies, a dummy variable for R&D stock equals zero and a dummy variable for IntSpill equals zero.

[a]The set of pre-sample means: Market Share, Employees, Tobin's Q, Sales, Assets, R&D stock, Patents stock and Citations stock.

Table A4

The effect of IntSpill and ExtSpill on Tobin's Q

Dependent variable: Log(Tobin's-Q); 9,454 observations

	(1)	(2)	(3)	(4)
R&D stock/Assets[a]	0.137* (0.025)	0.233* (0.041)	0.292* (0.048)	0.353* (0.056)
(R&D stock/Assets)		0.243* (0.045)	0.344* (0.061)	0.484* (0.091)
$(\text{R\&D stock/Assets})^2$		-0.013* (0.005)	-0.069* (0.026)	-0.184* (0.062)
$(\text{R\&D stock/Assets})^3$			0.005* (0.002)	0.029* (0.012)
$(\text{R\&D stock/Assets})^4$				-0.001* (0.001)
IntSpill x (R&D stock/Assets)[a]	0.044* (0.015)	0.041* (0.013)	0.059* (0.013)	0.063* (0.015)
IntSpill x (R&D stock/Assets)		0.087* (0.031)	0.188* (0.047)	0.266* (0.067)
IntSpill x $(\text{R\&D stock/Assets})^2$		-0.022* (0.016)	-0.119* (0.035)	-0.233* (0.070)
IntSpill x $(\text{R\&D stock/Assets})^3$			0.016* (0.005)	0.056* (0.021)
IntSpill x $(\text{R\&D stock/Assets})^4$				-0.003* (0.001)
ExtSpill x (R&D stock/Assets)a	-0.004* (0.002)	-0.005* (0.001)	-0.007* (0.002)	-0.008* (0.002)
ExtSpill x (R&D stock/Assets)		-0.012* (0.004)	-0.024* (0.006)	-0.0299* (0.012)
ExtSpill x (R&D stock/Assets)2		0.002* (0.001)	0.009* (0.003)	0.016* (0.009)
ExtSpill x (R&D stock/Assets)3			-0.001* (0.0002)	-0.002* (0.002)
ExtSpill x (R&D stock/Assets)4				0.0001* (0.0001)
Pre-sample means[b]	Yes	Yes	Yes	Yes

[a]Estimated marginal effects, evaluated at the mean. Standard errors are calculated using the Delta method.

Standard errors (in brackets) are robust to arbitrary heteroskedacity and serial correlation (Newey-West corrected). * denotes a significance level of 5 percent. All regressions include 78 two-digits industry dummies, 4 technology indicators, a complete set of year dummies, a dummy variable for R&D stock equals zero and a dummy variable for IntSpill equals zero.

[a]The set of pre-sample means: Market Share, Employees, Tobin's Q, Sales, Assets, R&D stock, Patents stock and Citations stock.

Table A5

The effect of IntSpill and ExtSpill on private returns to innovation: adding four-digit industry dummies

Dependent variable: Log(Tobin's-Q); 9,015 observations, 475 firms

	(1)	(2)	(3)	(4)
R&D stock/Assets[a]	0.197* (0.042)	0.342* (0.068)	0.199* (0.042)	0.348* (0.068)
R&D stock/Assets		0.366* (0.074)		0.373* (0.074)
$(\text{R\&D stock/Assets})^2$		-0.030* (0.009)		-0.031* (0.009)
IntSpill x (R&D stock/Assets)[a]	0.071* (0.021)	0.054* (0.015)	0.072* (0.019)	0.055* (0.014)
IntSpill x (R&D stock/Assets)		0.109* (0.041)		0.111* (0.039)
IntSpill x $(\text{R\&D stock/Assets})^2$		-0.021 (0.029)		-0.022 (0.029)
Externalized Spillovers x (R&D stock/Assets)[a]	-0.008* (0.002)	-0.005* (0.002)	-0.007* (0.002)	-0.005* (0.002)
ExtSpill x (R&D stock/Assets)		-0.014* (0.005)		-0.012* (0.005)
ExtSpill x (R&D stock/Assets)2		0.002* (0.0006)		0.002* (0.0006)
log(IntSpill)			0.005 (0.019)	0.005 (0.018)
log(ExtSpill)			-0.009 (0.014)	-0.011 (0.014)
Sales Growth	0.551* (0.055)	0.551* (0.055)	0.549* (0.054)	0.573* (0.054)
Pre-sample means[b]	Yes	Yes	Yes	Yes
Four-digit Industry effects	Yes	Yes	Yes	Yes
R^2	0.569	0.568	0.569	0.572

[a]Estimated marginal effects, evaluated at the mean. Standard errors are calculated using the Delta method. Standard errors (in brackets) are robust to arbitrary heteroskedacity and serial correlation (clustered at the firm level). * denotes a significance level of 5 percent.

All regressions include a complete set of year dummies, and a dummy for R&D stock equals zero and a dummy for Internalized Flows equal zero.

[b]The set of pre-sample means: Market Share, Employees, Tobin's Q, Sales, Assets, R&D stock, Patents stock and Citations stock.