**THE LAG STRUCTURE OF THE RELATIONSHIP BETWEEN PATENTING AND INTERNAL R&D REVISITED**

**Ning Wang**

Department of Organization and Strategy

Maastricht University

The Netherlands

Tel (31) 43-3884793

Fax (31) 43-3884893

N.Wang@maastrichtuniversity.nl

**John Hagedoorn**

Department of Organization and Strategy and UNU-MERIT

Maastricht University

The Netherlands

Tel (31) 43-3883823

Fax (31) 43-3884893

&

Department of Management

Royal Holloway, University of London

Egham, Surrey

United Kingdom

J.Hagedoorn@maastrichtuniversity.nl

*8 Apr, 2013 First Version*

*13 Nov, 2013 Revised*

*Research Policy*

**The Lag Structure of the Relationship between Patenting and Internal R&D Revisited**

**Abstract**

The principal purpose of this study is to revisit the classic research question of the lag structure of the patents-R&D relationship through an examination of the impact of internal R&D on firm patenting in the context of the global pharmaceutical industry during 1986–2000. Our empirical analysis, using both a multiplicative distributed lag model and a dynamic linear feedback model, differs from previous work that examines the patents-R&D relationship in three aspects. First, our estimation results exhibit direct evidence on lagged R&D effects, with the first lag (t-1) of R&D being significant in all distributed lag specifications. Second, a U-shaped lag structure of the patents-R&D relationship is found in most estimations of the multiplicative distributed lag model, which suggests a potential long-run effect of internal R&D investments on firm patenting. Finally, the results from the dynamic linear feedback model coincide with those from the multiplicative distributed lag model, indicating not only lag effects from recent R&D investments but also an overall long-run effect of internal R&D investments in the distant past on the knowledge production or innovation process of incumbent pharmaceutical firms.

**Keywords:**Patents;Internal R&D; Lag Structure; Multiplicative Distributed Lag Model; Dynamic Linear Feedback Model; Pharmaceutical Industry

1. **INTRODUCTION**

In the attempt to learn about gestation lags in knowledge production of in-house research and development (R&D) by firms, researchers have repeatedly examined the relationship between R&D expenditures and patents, which are taken as an output indicator of R&D (Bound *et al.*, 1984; Griliches, 1990)[[1]](#footnote-1). The question of interest is the lag structure of the patents-R&D relationship, studied by considering the number of patents applied for and granted to firms as a function of their current and lagged R&D expenditures. Pakes and Griliches (1984a) is probably the first attempt to look at the time shape of the distributed lag between patenting and internal R&D activity of firms. In their panel-data model (with a log-log functional form), Pakes and Griliches (1984a) found evidence of a lag truncation effect in the distributed lag of R&D on patents. The estimated coefficient on the last lag of R&D, with five lagged R&D terms in their model, was significantly higher than the coefficients of more recent R&D[[2]](#footnote-2).

 Hausman *et al*. (1984) and Hall *et al*. (1986) analyzed the same research question whether there is a lag in the relationship between patenting and R&D expenditures. Using a more appropriate functional form that reflected explicitly the non-negativity and discreteness of patent counts in the context of panel data, Hausman *et al*. (1984) found a U-shaped lag structure in the random-effects estimation but not in their conditional fixed-effects version. When they conditioned their estimates on the total number of patents received by a firm over the observed years, no coefficients except for the contemporaneous R&D were statistically significant either in the Poisson or negative binomial model. Hall *et al.* (1986) found similar results and concluded that there was very little direct evidence of anything but simultaneity in the year-to-year movement of patents and R&D expenditures, though an indirect analysis performed in their study suggested a possible distributed lag structure.

 The consistency of the previously described panel-data models[[3]](#footnote-3) rests on the assumption that patents are an indicator of the output or ‘success’ of R&D rather than the input of R&D (Hall *et al.*, 1986). However, as the patent application tends to occur relatively early in the life of a research project and the bulk of R&D expenditures often occur after the application is made, new patents virtually generate the need for future R&D expenditures (Griliches, 1990; Hall *et al.*, 1986). Therefore, R&D expenditures should be seen as a predetermined variable instead of a strictly exogenous one. Such concern in the relationship of patents to internal R&D activity was first addressed by Hall *et al.* (1986) using a simple version of a Granger causality test, with a view to testing if past success in patenting leads to an increase in a firm’s future R&D investments[[4]](#footnote-4). Montalvo (1997) applied a quasi-differenced generalized method of moments (GMM) estimator to the analysis of the patents-R&D relationship so as to obtain consistent estimates in the presence of predetermined regressors, i.e., R&D expenditures[[5]](#footnote-5). The results turned out to be somewhat inconclusive as well: the estimated coefficient on contemporaneous R&D was not statistically significant while the first lag of R&D had a significant effect on patents.

 Blundell *et al.* (2002) extended the quasi-differenced GMM estimation with an application to a dynamic linear feedback model and proposed an alternative estimator, the pre-sample mean (PSM) estimator, based upon pre-sample information on the dependent variable. In their application to the analysis of the patents-R&D relationship, the results for the dynamic linear feedback model from the PSM estimator indicated a much lower depreciation rate of internal R&D investments—a potential long-run effect of in-house R&D on firm patenting—than that implied by the results from the multiplicative distributed lag model in prior literature. A recent study on the patents-R&D relationship by Gurmu and Pérez-Sebastián (2008) reported lagged R&D effects that were moderately higher than those previously found, but the lag effects on patents were identified only for more recent R&D.

So far the earlier work in this area, as aforementioned, hasinvestigated the relationship between patenting and internal R&D activity of firms for the U.S. manufacturing sector during the 1970’s (Blundell *et al.*, 2002; Hall *et al*., 1986; Hausman *et al*., 1984; Montalvo, 1997; Pakes and Griliches, 1984a) and over the 1980’s (Gurmu and Pérez-Sebastián, 2008). Our study aims to revisit this classic research question regarding the lag structure of the patents-R&D relationship by applying recently developed estimation techniques on firm-level panel data for the global pharmaceutical industry from 1986 to 2000. Prior research suggests that the relationship between patenting and internal R&D activity of firms differs across industries (Griliches, 1990; Hall *et al.*, 1986). We attempt to address this concern by taking a closer look at the lag structure of the patents-R&D relationship within one industry. We focus on the global pharmaceutical industry for two main reasons. First, the pharmaceutical industry as a high-technology sector is characterized by high-levels of patenting propensity and R&D intensity. Previous studies demonstrate that patenting activity is an important source of technological advantage in the pharmaceutical industry (Henderson and Cockburn, 1994; Levin *et al.*, 1987). In addition, recent figures show that pharmaceutical firms invest as much as five times more in R&D, relative to their sales, than the average U.S. manufacturing firm (CBO Study, 2006)[[6]](#footnote-6). Second, empirical evidence clearly indicates that the proportion of research (‘R’) in R&D expenditures is the main contributor to patents, whereas the bulk of development (‘D’) costs lead more to products and processes (Czarnitzki *et al.*, 2009)[[7]](#footnote-7). Given that the ‘D’ part of R&D expenditures, relative to ‘R’, is mostly the larger one, the estimated patents-R&D elasticity would be biased downwards when development costs are of minor relevance for patent production (Czarnitzki *et al.*, 2009; Griliches, 1990). As the pharmaceutical industry is actually among the most research-intensive sectors with a very large share of ‘R’ (CBO Study, 2006), studying the pharmaceutical industry may alleviate the above-mentioned problem ~~that the ‘relevant’ R&D is measured with error (Griliches, 1990) in the patents-R&D relationship~~.

 Our empirical analysis, using both a multiplicative distributed lag model and a dynamic linear feedback model, differs from previous work that examines the patents-R&D relationship in three aspects. First, our estimation results exhibit direct evidence on lagged R&D effects, with the first lag (t-1) of R&D being significant in all distributed lag specifications. Evidence for the contribution of the first lag of R&D to the current year’s patent counts is of more than 25% of the total R&D elasticity. Second, a U-shaped lag structure of the patents-R&D relationship is found in most estimations of the multiplicative distributed lag model. This finding suggests a potential long-run effect of internal R&D investments on firm patenting. Finally, the estimation results from the dynamic linear feedback model coincide with those from the multiplicative distributed lag model, indicating not only lag effects from recent R&D investments but also an overall long-run effect of internal R&D investments in the distant past in the knowledge production or innovation process of incumbent pharmaceutical firms.

 The rest of this paper is structured as follows. In section 2, we provide the theoretical background to the expected lag effects of internal R&D investments by firms, lags effects that may last during the long run of the knowledge production or innovation process. In that section, we also formulate our main hypothesis. Section 3 describes the derivation of the data set and looks at the properties of the various variables. Section 4 proceeds by presenting the two count panel data models underlying our empirical analysis—the multiplicative distributed lag model and the dynamic linear feedback model—and their associated estimation techniques. The empirical results are then reported in Section 5. Section 6 summarizes our main results and their implications and it discusses some possible future lines of work.

1. **THEORETICAL BACKGROUND AND HYPOTHESIS**

Knowledge production or innovation is often described as a recombinant process that involves combining existing bits of knowledge in the creation of new knowledge (Fleming, 2001; Kogut and Zander, 1992). A firm’s existing bits of knowledge or knowledge stock is built up over time through its previous investments in R&D (Hall *et al.*, 1986). In the context of the relationship between patents and R&D expenditures, the annual R&D expenditures are considered to be investments which add to a firm’s knowledge stock, and patents are taken as an output indicator of new knowledge creation or innovation (Hall *et al.*, 1986; Hausman *et al.*, 1984). As argued by Scotchmer (1991), the innovation process is of a cumulative nature such that early innovations (i.e., knowledge stock) provide a boost or a technological foundation for later innovations (i.e., new knowledge)[[8]](#footnote-8). This forms a theoretical rationale underlying the lag effects of R&D investments in the sense of contributions to firm innovativeness.

 A firm’s knowledge stock usually depreciates as time passes (Hall *et al.*, 1986) since knowledge tends to become obsolete and it no longer matches the demands of its current environment (Eisenhardt, 1989; Helfat and Raubitschek, 2000). In this respect, more recently created knowledge is often considered more valuable for knowledge production or innovation because of organization-environment fit (Sørensen and Stuart, 2000), capability building in emerging areas (McGrath, 1999), temporal local search (Helfat, 1994; Martin and Mitchell, 1998), and cognitive and institutional pressures (Martins and Kambil, 1999; Nerkar, 2003). In other words, in the case of the patents-R&D relationship, recent R&D investments are expected to have a greater impact on firm patenting, while the contribution of older R&D has become less valuable with the passage of time.

 However, research also suggests that recent knowledge may not necessarily be representative of the technologically superior or optimal solution that has emerged (Abrahamson, 1996). In contrast, a firm’s older knowledge spread across long time spans, accruing from its R&D investments in the distant past, may also be influential with respect to new knowledge creation or innovation (Garud and Nayyar, 1994; Nerkar, 2003). For example, incumbent pharmaceutical firms often review their previously discarded experimental compounds, some of which failed in clinical trials as long as twenty years ago, hoping that an old compound intended for one treatment may be useful in treating something different nowadays (Simons, 2006). The rationale for the important role of older knowledge is that it is characterized by increased reliability and legitimacy, as older knowledge is usually better tested and understood by firms than recently created knowledge, thereby decreasing the chances of costly errors and increasing the likelihood of successful innovation (Katila, 2002; March, 1991). The far-reaching influence of older knowledge implies that even long past R&D investments by firms can still play a crucial part in their knowledge production or innovation, or put differently, the lag effects produced by internal R&D investments may last over a long time horizon.

 Such potential long-run effect of in-house R&D is consistent with a real options logic for managing R&D investment strategies of firms. As stated by Myers (1984) “ … the value of R&D is almost all option value …”. ~~R&D investments can be considered to be real options that convey the right for firms to preserve their decision rights in the future in their investment choices (Garud and Nayyar, 1994; McGrath, 1997).~~ Comment JH: similar sentence appears two sentences further down. For firms operating in a rapidly changing environment, it is important to constantly direct their R&D efforts towards new technological areas that are often characterized by high levels of technical uncertainty.[[9]](#footnote-9) In the face of technical uncertainty, it is in the best interests of firms to invest immediately, otherwise they might be subject to a discounting penalty or even the risk of lock-out (McGrath, 1997). Such investments in R&D are viewed as real options that create value by providing firms with the luxury of waiting and watching as more information about uncertainty is revealed (McGrath, 1997; Nerkar, 2003). Specifically, some knowledge derived from past R&D investments may demonstrate a great potential value at a future date when conditions favor its (re)use, for instance, through coevolution of complementary knowledge, markets, institutions, or standards that are necessary for employing useful but untapped knowledge (Garud and Nayyar, 1994; Nerkar, 2003). To the extent that it might take a very long time for these favorable conditions to emerge, there are time lags for firms’ past knowledge to become conducive for future knowledge creation or innovation. Furthermore, in addition to waiting and watching after the initial investment in R&D, i.e., the opened option, firms may have strong incentives to persist in making investments in the same area (McGrath and Nerkar, 2004). This is due to increases in absorptive capacity that, with the cumulative experience through previous track record of investments, makes the assimilation of subsequent knowledge in the same area easier, faster, and less expensive (Cohen and Levinthal, 1990; McGrath and Nerkar, 2004). Such sequential investing in internal R&D leads to a cumulative, path-dependent knowledge production or innovation process, which makes the very early option investments in R&D valuable or even vital for a firm’s success in a new technological area. Hence, from the perspective of real options reasoning, early R&D investments by firms may generate a long-lasting impact on future knowledge creation or innovation.

 The above indicates that a firm’s internal R&D investments give rise to, not only lag effects in the short run that arise from more recent R&D, but also a potential long-run effect resulting from its distant R&D history in the knowledge production or innovation process. We therefore propose the following main hypothesis:

*Hypothesis: Taking patents as an output indicator of firms’ innovativeness, lag effects of firms’ internal R&D investments are expected to impact their long run knowledge production or innovation process.*

1. **DATA AND VARIABLE CONSTRUCTION**

The research setting for this study is the global pharmaceutical industry, which is on the verge of profound mutations as a new biotechnology-based technological regime has emerged. More in particular, we will focus on incumbent pharmaceutical firms which attempt to build up innovative capabilities within this new technological regime. Incumbent pharmaceuticals are defined as pharmaceutical firms that were in existence prior to the emergence of biotechnology. These companies, such as Bayer, Hoffmann-La Roche, Merck, and Pfizer, are generally mature and very large firms that dominated the industry since the 1940s. In the mid-1970s, new biotechnology brought along significant scientific and technological breakthroughs in genetic engineering (recombinant DNA, 1973) and hybridization (monoclonal antibodies, 1975). These advances virtually serve as a radical process innovation for incumbent pharmaceutical companies in the way new drugs are discovered and developed (Pisano, 1997).

 To mitigate a potential survivor bias, we started with a comprehensive set of incumbent pharmaceutical firms alive as of 1986 based on various industry sources[[10]](#footnote-10). Through this process, we identified 89 incumbent pharmaceutical firms, defined as pharmaceutical firms that were founded prior to the emergence of biotechnology in the mid-1970s. A further scrutiny criterion for the sample firms is based on the absence of large jumps during the time period of 1986–2000. Following Hall *et al.* (1986), a jump is defined as an increase in total assets or employment of more than 100 per cent or a decrease of more than 50 per cent[[11]](#footnote-11). This test was not applied unless the change in total assets was greater than two million dollars or the change in employment was greater than 500 employees. Our study ultimately ended up with a slightly unbalanced panel sample of 85 firms covering the 15-year period from 1986 to 2000, giving 1,186 firm-year observations.

 An important source of data for our study is the Technology Profile Report maintained by the U.S. Patent and Trademark Office (USPTO). All patent data used in this study are based on successful patent applications, or granted patents. The application date of a patent is only reported when the patent is actually granted. On average it may take three years for the USPTO to grant a patent after it was originally applied for by an inventing firm (Biotechnology Industry Organization, 2008). In contrast, there is essentially no lag time between the completed invention and the patent application date, which is on average no more than three months (Darby and Zucker, 2003). Thus, the application date of a granted patent is closely tied to the timing of the new knowledge creation and should be used as the relevant time placer for patents (Hall *et al.*, 2001; Trajtenberg, 1990). Accordingly, we assigned a granted patent to the year in which the patent was originally applied for. For instance, a patent applied for in 1986 but granted in 1989 is considered a 1986 patent. We use USPTO patent data for both US and non-US firms. Although US patent data could imply a bias against non-US firms, it is mentioned in the literature that, given the importance of the US market, the patent protection offered by US authorities, and the level of technological sophistication of the US market, it is inevitable that non-US firms, in particular leading, large firms file their patents in the USA (Hagedoorn and Cloodt, 2003; Patel and Pavitt, 1991). Furthermore, as suggested by Ahuja and Katila (2001), to maintain a certain level of consistency, reliability and comparability choosing one patenting system is preferred to using several patenting systems across nations.

 Financial figures and employment data were obtained from Compustat and Datastream (Thomson Financial). For all firms, financial data were converted to U.S. dollars and are inflation-adjusted in millions of year 2000 dollars.

**Dependent Variable**

The dependent variable for this study is *the number of patents* applied for by an incumbent pharmaceutical firm in a given year that are eventually granted, which is taken as an output indicator of advances in knowledge, or more in general, firm innovativeness. As displayed in Table 1, the frequency distribution of the number of patents is characterized by heavy upper tails with a range of 0 to 800, a mean value of about 94, and a relatively low median value of 45. The first and third quantile values are 14 and 116 patents, respectively. Over the whole study period, the proportion of zero patents is only 1%, while the fraction with at least 100 patents is 29%. Compared to previous studies (See e.g., Cincera, 1997; Gurmu and Pérez-Sebastián, 2008; Hall et *al.*, 1986; Hausman *et al*., 1984), the variance of the number of patents is slightly larger than the mean, indicating the presence of a moderate overdispersion in our patent data[[12]](#footnote-12).

*Insert Table 1 here*

**Independent Variable**

The main explanatory variables of interest are the natural logarithms of current and lagged values of internal R&D investments. Wemeasured an incumbent pharmaceutical firm’s internal R&D investments by its annual *R&D expenditures*. The annual R&D expenditures are considered to be investments that add to a firm’s stock of knowledge, whose value decays over time such that the contribution of older R&D investment becomes less valuable as time passes (Hall *et al.*, 1986). Table 1 shows that incumbent pharmaceuticals are R&D-intensive firms, their average R&D expenditures amounting to approximately 505 million dollars per year. From Table 2 we can see that R&D expenditures by incumbent pharmaceutical firms are highly correlated over time, with the first order autocorrelation of about 0.99 and the correlation among R&D regressors all above 0.95. We will discuss this issue of a highly persistent R&D series at greater length in the forthcoming sections. R&D expenditures are inflation-adjusted in millions of year 2000 dollars.

*Insert Table 2 here*

**Control Variables**

The logarithm of *pre-sample mean* *of patents* is included to proxy for the unobserved heterogeneity in firm knowledge production (Blundell *et al.*, 1995; Blundell *et al.*, 2002). We used the average number of patents by a pharmaceutical firm in the period from 1981 to 1985, i.e., a five-year period of pre-sample history, to construct the pre-sample mean variable.

 Following previous studies (Gurmu and Pérez-Sebastián, 2008; Hall *et al*., 1986; Hausman *et al*., 1984), we used a time-constant regressor, i.e., the logarithm of book value of equity in year 2000 in millions of dollars, as a measure for firm size[[13]](#footnote-13). For those firms that fail to survive till year 2000, the end year value of equity is used and inflation-adjusted.

 To control for country-specific institutional configurations that are significant in shaping firms’ propensities to patent (Jaffe and Trajtenberg, 1999), we included two indicator variables based upon the location of company headquarters. One indicator variable is coded as 1 if a pharmaceutical firm is located in the United States (*1 = U.S. Firm*), and the other indicator is coded as 1 if a pharmaceutical firm is headquartered in Europe (*1 = European Firm*), with a Japanese location as the reference category. The global nature of our data set is reflected in the fact that 33% of the firms are U.S. based, 30% are European based, with the remaining 37% headquartered in Japan.

 In contrast to previous studies (Hall *et al*., 1986; Hausman *et al*., 1984; Gurmu and Pérez-Sebastián, 2008) where a scientific sector dummy was used to distinguish R&D-intensive industries from other manufacturing industries, our study focuses on one high-tech industry in the scientific sector, the global pharmaceutical industry. Since the pharmaceutical industry is made up of both specialized pharmaceutical firms and more diversified, mainly chemical, conglomerates, we controlled for a pharmaceutical firm’s degree of diversification by coding the firm as 1 if it is a specialized company (*1 = Pharma Firm*) and 0 otherwise. Specialized pharmaceutical firms are those active in SIC 2834 (Pharmaceutical Preparations Manufacturing), whereas a conglomerate might engage, for instance, in both SIC 2834 and SIC 2890 (Chemical Products Manufacturing). About one half of the firms in the sample are fully specialized (46%).

 Over time, patenting frequency may increase or decrease for all firms. To control for such time-varying effects, we included *year fixed effects* for each year, with 2000 being the reference year.

1. **ESTIMATION**

As the dependent variable of this study, the number of patents applied for and received by firms, is a count variable taking only non-negative integer values, we employ various count panel data models—the multiplicative distributed lag model and the dynamic linear feedback model—to investigate the relationship between patenting and internal R&D activity of firms in the global pharmaceutical industry. For simplicity, all estimation methods to be used in our empirical analysis will only be sketched here. The relevant models and the associated moment conditions to be solved are described in Appendix B (See GMM І and PSM І for the estimation of Model (1), the multiplicative distributed lag model, and GMM II and PSM II for estimating Model (2), the dynamic linear feedback model).

 **4.1 Multiplicative Distributed Lag Model**

The conditional mean for a standard multiplicative distributed lag model is specified as

(1)

where are the number of patents applied for and received by firm i in period t, denotes annual R&D expenditures, is a vector of observable firm-specific effects (such as firm size, firm nationality, and Pharma firm)[[14]](#footnote-14), is a vector of time-specific variables (i.e., year fixed effects), and represents unobserved individual fixed effects, which are commonly modeled multiplicatively in count panel data models.

 The basic reference for estimating the above model is the Poisson and negative binomial estimations (Hausman *et al*., 1984). Both random- and fixed-effects specifications are used in this study to control for unobserved individual fixed effects, , in the Poisson and negative binomial estimations. The problem with the random (uncorrelated) effects formulation is that it is inconsistent when is correlated with the regressors of interest[[15]](#footnote-15). Conditional fixed-effects specification, however, allows for such correlations (Hall *et al.*, 1986; Hausman *et al*., 1984). It is also important to note that, on account of the overdispersed nature of patent counts, the negative binomial estimation provides a better fit for our data than the Poisson estimation since the assumption of the equality of conditional mean and conditional variance may not hold (Hall *et al.*, 1986; Hausman *et al*., 1984).

 The consistency of the standard estimators presented above relies on the strict exogeneity of the explanatory variables. As previously discussed, new patents may depend on additional R&D expenditures in the future for their full development, the series of current and lagged R&D expenditures is therefore weakly exogenous or predetermined in our estimation models. As an alternative to the standard estimators, Chamberlain (1992) and Wooldridge (1997) have proposed a quasi-differenced GMM estimator that relaxes the strict exogeneity assumption. Specifically, they provide a transformation strategy that eliminates the unobserved fixed effects and then use valid moment conditions to obtain consistent estimation. In this study, we follow Wooldridge’s (1997) quasi-differencing transformation and use the instruments) as below, which require a restricted serial correlation of the residuals[[16]](#footnote-16).

 On the other hand, the GMM estimator is subject to some problems, namely small sample bias and imprecision when the series are highly persistent, as the instruments are then weak predictors of the endogenous changes (Blundell *et al*, 1995; Blundell *et al.*, 2002). As an alternative, Blundell *et al.* (2002) propose a pre-sample mean (PSM) estimator that replaces the unobserved fixed effect by the pre-sample mean of the dependent variable. In view of the highly persistent R&D series of this study, as will be exhibited in Table 2, we used the PSM estimator as well.

 **4.2 Dynamic Linear Feedback Model**

An additional complication to the highly persistent R&D series is that the current and lagged R&D terms in the multiplicative distributed lag model [i.e., Model (1)] are collinear. To solve this problem, Blundell *et al.* (2002) develop a dynamic linear feedback model by explicitly introducing the dynamics of the count process in panel data. In contrast with Model (1), in the dynamic linear feedback model, the lagged dependent variable enters the conditional mean specification linearly, which corresponds to the following[[17]](#footnote-17)

 (2)

where estimates the depreciation factor, and are the long-run and short-run elasticities of patents with respect to internal R&D, respectively (See Appendix B for an economic interpretation of the dynamic linear feedback model).

 Following Blundell *et al.* (2002), we use two approaches to obtaining estimates for the dynamic linear feedback model [i.e., Model (2)]. The first is the GMM estimator, based on a quasi-differencing transformation proposed by Wooldridge (1997) and a choice of the instruments as follows:

A further estimation of Model (2) using the PSM estimator, as discussed earlier, includes pre-sample information on the number of patents to control for individual fixed effects. Both estimators are consistent in the presence of predetermined regressors and correlated fixed effects.

1. **RESULTS**

The main findings for Model (1), in which five lagged R&D terms are included[[18]](#footnote-18), and Model (2) are presented in Table 3. Columns 1 through 4 give results from the standard Poisson and negative binomial estimations of Model (1). The coefficients on contemporaneous and the first lag (t - 1) of R&D are positive and significant across all the Poisson and negative binomial regression models. Moreover, a consistent pattern of a U-shaped lag structure is found, with the first (t) and last (t - 5) coefficients being larger than those in the middle, for the Poisson and negative binomial models in both the random- and fixed-effects specifications.

 The significant, positive, and large coefficient on the last lag (t-5) of R&D could be due to the correlation of the last R&D variable with earlier left-out R&D. If this is the case, a U-shaped lag structure is solid evidence for a lag truncation effect in the distributed lag of R&D on patents (Hall *et al*., 1986; Pakes and Griliches, 1984a). To explore this idea further, we performed autoregressive regressions on the R&D series. As shown in Table A1 (see Appendix A), an AR (2) process can be accepted at 5% significance level. Nonetheless, the first lag coefficients are close to one and the second lag coefficients are relatively small, which implies that it is difficult to reject a random walk process. Hence, the R&D series of this study conforms to a random walk or an AR (2) process and suggests the existence of a lag truncation effect.

 On the other hand, the fairly large coefficient on the last lagged R&D may also be caused by correlated fixed effects, i.e., the correlation between the permanent patenting propensity of firms and their investments in internal R&D (Hall *et al*., 1986; Pakes and Griliches, 1984a). We then used fixed-effects models that allow for such correlations, and a U-shaped lag structure was identified as well in the fixed effects formulation. This finding suggests that correlated fixed effects do not bias the estimated results and there is indeed a lag truncation effect in the distributed lag relationship between patenting and R&D expenditures. By contrast, previous work is inconclusive about the lag structure of the patents-R&D relationship. Hausman *et al*. (1984) and Hall *et al*. (1986) found a U-shaped lag structure for the Poisson and negative binomial random-effects models, but only a contemporaneous relationship was identified in their conditional fixed-effects models. In the recent study by Gurmu and Pérez-Sebastián (2008), a U-shaped lag structure was found for the random- and fixed-effects Poisson models but no evidence of it in the case of negative binomial models.

*Insert Table 3 here*

 Results for Model (1) using the less constrained GMM and PSM estimators, which relax the strict exogeneity assumption, are reported in columns GMM І and PSM І, respectively (see Table 3 Continued)[[19]](#footnote-19). In the GMM estimation, a Sargan Difference test is used to check for overidentifying restrictions, and m1 and m2 are tests for first- and second-order serial correlation. As shown in column GMM І, the Sargan test does not reject the null hypothesis that the additional moment conditions are valid. In addition, the first- and second-order serial correlation tests are consistent with the serially uncorrelated error term (i.e., significant m1 and insignificant m2). The specification tests thus do not indicate misspecification in GMM І. Comparison of the results from GMM І to those from the Poisson and negative binomial estimations shows that there are still strong contemporaneous and lag effects of R&D, with significant positive coefficients on t, t-1, and t-5. However, a negative and significant effect is found at the second lag (t-2) of R&D. Although this is not as expected, similar cases also occur in previous studies (see Gurmu and Pérez-Sebastián, 2008; Montalvo, 1997)[[20]](#footnote-20). The puzzling estimated form of the lag structure in GMM І possibly arises from a problem of weak instruments, as the R&D series in our data set is highly persistent over time (the first order autocorrelation is about 0.99 as displayed in Table 2) and the instruments are only weakly correlated with the endogenous variables in the differenced model (Bond and Windmeijer, 2005; Blundell *et al*, 1995; Blundell *et al.*, 2002).

 In column PSM І, results are given from the PSM estimator for the estimation of Model (1). The pre-sample mean of patents is positive and statistically significant, indicating that it is important to control for the unobserved differences in the capabilities for knowledge production with which firms entered the sample. With regard to the influence of internal R&D, both the contemporaneous and first lagged R&D terms retain their signs, though with reduced level of significance. The significant effect of the last lag of R&D, however, no longer appears in the PSM І version. The explanation lies in the fact that the information contained in the pre-sample mean becomes close to that reflected by the R&D series when the R&D process is highly persistent. Thus, the separate estimation of the parameters on R&D and the pre-sample mean is problematic due to multicollinearity (Blundell *et al.*, 2002). Since the variance of the fixed effect (proxied by the pre-sample mean) is relatively large, the parameter on the pre-sample mean variable will get a larger weight, leading to a downward bias in the parameters for R&D series[[21]](#footnote-21).

 To give a brief summary of Model (1), based on the various estimations presented above, in all cases the results suggest that there are lag effects of internal R&D as well as a significant contemporaneous relationship between patents and R&D. Consistent with prior literature, the impact of contemporaneous R&D on patents is rather strong. Nonetheless, our study differs substantively from previous work in that the estimation results exhibit direct evidence on lagged R&D effects, with the first lag of R&D being significant in all distributed lag specifications. More importantly, a U-shaped lag structure of the patents-R&D relationship is identified across the various estimations of Model (1) (except for PSM І), suggesting a potential long-run effect of internal R&D investments on firm patenting.

 We now turn to the estimation of Model (2), the dynamic linear feedback model, which essentially allows for the dynamics of the count process itself. Results for Model (2) using the GMM and PSM estimators are reported in columns GMM II and PSM II, respectively (see Table 3 Continued). As seen in GMM II, the specification tests provide no evidence against the model. The estimated coefficients on the lagged dependent variable and R&D equal to 0.21 ( ) and 0.73 (. This implies that the contribution of internal R&D investments depreciates exponentially at the rate of approximately = 79%, with the short-run elasticity of patents with respect to R&D of about = 0.57 and the long-run elasticity of about = 0.73. These results provide an indication of lag effects of R&D expenditures on patents, though at a relatively fast moving process. The suspicion is still that, as in the case of GMM І, the GMM estimator is subject to small sample bias and the weak instrument problem, particularly when R&D series is highly persistent over time (Blundell *et al*, 1995; Blundell *et al.*, 2002).

 The estimates from PSM II show a rather different picture, with a fairly low depreciation rate of approximately 10% and the long-run elasticity of patents with respect to R&D of about 0.50, which is substantially larger than the short-run elasticity, 0.05. Therefore, PSM II implies a much slower moving process and accordingly an overall long-run effect of in-house R&D on firm patenting. In this vein, the estimation results from Model (2), the dynamic linear feedback model, coincide with those from Model (1), the multiplicative distributed lag model, indicating not only lag effects but also a potential long-run effect of internal R&D investments in a firm’s knowledge production and innovation process.

*Insert Table 3 Continued here*

1. **DISCUSSION AND CONCLUSIONS**

The principal purpose of this study is to revisit the classic research question of the lag structure of the patents-R&D relationship through an examination of the impact of internal R&D on firm patenting in the context of the global pharmaceutical industry during 1986–2000[[22]](#footnote-22). Our empirical analysis shows that, although our results are somewhat sensitive to different estimation methods, the total elasticity of patenting with respect to R&D varies from 0.3 to 0.9, indicating decreasing returns to scale in internal R&D. Consistent with prior research, the impact of contemporaneous R&D on patents is rather strong, accounting for over 36% of the overall R&D elasticity. However, our study differs from previous work that examines the patents-R&D relationship in three aspects. First, our estimation results exhibit direct evidence on lagged R&D effects, with the first lag (t-1) of R&D being significant in all distributed lag specifications. Evidence for the contribution of the first lag of R&D to the current year’s patent counts is of more than 25% of the total R&D elasticity. Second, a U-shaped lag structure of the patents-R&D relationship is identified across the various estimations of the multiplicative distributed lag model (except for PSM І). This finding suggests a potential long-run effect of internal R&D investments on firm patenting. Finally, the estimation results for the dynamic linear feedback model, especially from the PSM estimator, coincide with those from the multiplicative distributed lag model, with a fairly low depreciation rate of approximately 10% of internal R&D investments over time and accordingly an overall long-run effect of internal R&D on firm patenting. In short, our empirical findings, from both the multiplicative distributed lag model and the dynamic linear feedback model, provide evidence for not only lag effects from recent R&D investments but, in support of our main hypothesis, also a potential long-run effect of internal R&D investments on the distant past in the knowledge production or innovation process of incumbent pharmaceutical firms.

 To summarize, our results illustrate a cumulative knowledge production or innovation process of incumbent pharmaceutical firms, with current patent production being historically dependent on past investments in internal R&D. These results are consistent with prior research which suggests that not only recent knowledge is important for new knowledge creation, but that older knowledge in knowledge stock accruing from long past R&D investments may also be valuable for firm innovativeness (Katila, 2002; Nerkar, 2003). This finding echoes the real options logic for managing R&D investment strategies of firms (McGrath, 1997; McGrath and Nerkar, 2004), according to which early R&D investments by firms in a new area with technical uncertainty may be influential in their knowledge production or innovation process over a long-term horizon. After the initial investment in R&D, firms can choose to wait until conditions favor the (re)use of the old untapped knowledge (McGrath, 1997; Nerkar, 2003). Alternatively, firms may opt for making further R&D investments in the same area in a sequential way, which enhances the value of the initial R&D investment and makes it valuable for future innovation.

 From a managerial perspective, our study offers important insights into the cumulative knowledge production or innovation process of firms. In addition to emphasizing recent investments in internal R&D to stay abreast of the latest, cutting edge technologies, decision-makers in firms should also adopt a long-term perspective for organizing R&D investment strategies. Through periodically reviewing and recombining the old, useful but under-utilized knowledge, firms can increase their creation of new knowledge (Garud and Nayyar, 1994; Nerkar, 2003). However, it is worth to note that old knowledge tends to be lost over time due to lack of adequate organizational memory, inaccurate recording, and turnover in R&D personnel (Argote, 1999). As a consequence, to effectively transfer knowledge across time, firms need to build up their ‘transformative capacity’, which pertains to the choice of knowledge for future use, its maintenance over time, and the reactivation and synthesis of such knowledge when required (Garud and Nayyar, 1994). In this way, by making better use of old knowledge and actively maintaining stock of knowledge for future use, firms can enhance their returns from internal R&D investments.

 Our study is subject to some limitations which suggest avenues for future research. The restriction of the sample to a single industrial context raises the question of the generalizability of our findings, thus further efforts could be made to conduct research in other high-technology industries. In addition, as the proportion of research (‘R’) in R&D expenditures is the main contributor to patents (Czarnitzki *et al.*, 2009; Griliches, 1990), future work directed towards a fine-grained analysis of disaggreating ‘R’ and ‘D’ would contribute to a better understanding of the overall lag structure of the patents-R&D relationship.

**REFERENCES**

Abrahamson, E. 1996. Management fashion. *Academy of Management Review* **21**: 254–286.

Ahuja, G., Katila, R., 2001. Technological acquisition and the innovative performance of acquiring firms: A longitudinal study. *Strategic Management Journal* **22**: 197–220.

Blundell, R., Griffith, R., Reenen, J.V. 1995. Dynamic count data models of technological innovation. *Economic Journal* **105**: 333–344.

Blundell, R., Griffith, R., Windmeijer, F. 2002. Individual effects and dynamics in count data models. *Journal of Econometrics* **108**: 113–131.

Bond, S., Windmeijer, F., 2005. Reliable inference for GMM estimators? Finite sample properties of alternative test procedures in linear panel data models. *Econometric Reviews* **24**: 1–37.

Bound, J., Cummins, C., Griliches, Z., Hall, B.H., Jaffe, A. 1984. Who does R&D and who patents? In: Griliches, Z. (ed.). *R&D, Patents, and Productivity*. Chicago: University of Chicago Press.

Chamberlain, G. 1992. Comment: sequential moment restrictions in panel data. *Journal of Business & Economic Statistics* **10**: 20–26.

Cincera, M. 1997. Patents, R&D, and technological spillovers at the firm level: Some evidence from econometric count models for panel data. *Journal of Applied Econometrics* **12**: 265–280.

Cohen, W.M., Levinthal, D.A. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* **35**: 128–152.

Crépon, B., Duguet, E. 1997. Estimating the innovation function from the patent numbers: GMM on count panel data. *Journal of Applied Econometrics* **12**: 243–264.

Czarnitzki, D., Kraft, K., Thorwarth, S. 2009. The knowledge production of ‘R’ and ‘D’. *Economics Letters* **105**: 141–143.

Darby, M.R., Zucker, L.G. 2003. Growing by leaps and inches: Creative destruction, real cost reduction, and inching up. *Economic Inquiry* **41**: 1–19.

Dixit, A., Pindyck, R. 1994. *Investment under Uncertainty*. Princeton, NJ: Princeton University Press.

Eisenhardt, K. 1989. Making fast strategic decisions in high-velocity environments. *Academy of Management Journal* **32**: 543–576.

Fleming, L. 2001. Recombinant uncertainty in technological search. *Management Science* **47**: 117–132.

Garud, R., Nayyar, P.R. 1994. Transformative capacity: Continual structuring by intertemporal technology transfer. *Strategic Management Journal* **15**: 365–385.

Griliches, Z. 1990. Patent statistics as economic indicators: A survey. *Journal of Economic Literature* **28**: 1661–1707.

Guo, J.Q., Trivedi, P.K. 2002. Flexible parametric models for long-tailed patent count distributions. *Oxford Bulletin of Economics and Statistics* **63**: 63–82.

Gurmu, S., Pérez-Sebastián, F. 2008. Patents, R&D, and lag effects: Evidence from flexible methods for count panel data on manufacturing firms. *Empirical Economics* **35**: 507–526.

Hagedoorn, J., Cloodt, M., 2003. Measuring innovative performance: Is there an advantage in using multiple indicators? *Research Policy* **32**: 1365–1379.

Hall, A.J., Sivamohan, M.V.K., Clark, N., Taylor, S., Bockett, G. 2001. Why research partnerships really matter: Innovation theory, institutional arrangements and implications for the developing new technology for the poor. *World Development* **29**: 783–797.

Hall, B.H., Hausman, J., Griliches, Z. 1986. Patents and R&D: Is there a lag? *International Economic Review* **27**: 265–283.

Hausman, J., Hall, B.H., Griliches, Z. 1984. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica* **52**: 909–938.

Helfat, C.E. 1994. Evolutionary trajectories in petroleum firm R&D. *Management Science* **40**: 1720–1747.

Helfat, C.E., Raubitschek, R. S. 2000. Product sequencing: Co-evolution of knowledge, capabilities and products. *Strategic Management Journal* **21**: 961–979.

Henderson, R., Cockburn, I. 1994. Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic Management Journal* **15**: 63–84.

Jaffe, A.B., Trajtenberg, M. 1999. International knowledge flows: Evidence from patent citations. *Economics of Innovation & New Technology* **8**: 105–136.

Katila, R. 2002. New product search over time: Past ideas in their prime? *Academy of Management Journal* **45**: 995–1010.

Kogut, B., Zander, U. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* **3**: 383–397.

Levin, R.C., Klevorick, A.K., Nelson, R.R., Winter, S.G. 1987. Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity* **3**: 783–831.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science* **2**: 71–87.

Martin, X., Mitchell, W. 1998. The influence of local search and performance heuristics on new design introduction in a new product market. *Research Policy* **26**: 753–771.

Martins, L. L., Kambil, A. 1999. Looking back and thinking ahead: Effects of prior success on managers’ interpretations of new information technologies. *Academy of Management Journal* **42**: 652–661.

McGrath, R.G. 1997. A real options logic for initiating technology positioning investments. *Academy of Management Review* **22**: 974–996.

McGrath, R.G. 1999. Falling forward: Real options reasoning and entrepreneurial failure. *Academy of Management Review* **24**: 13–30.

McGrath, R.G., Nerkar, A. 2004. Real options reasoning and a new look at the R&D investment strategies of pharmaceutical firms. *Strategic Management Journal* **25**: 1–21.

Montalvo, J.G. 1997. GMM estimation of count-panel-data models with fixed effects and predetermined instruments. *Journal of Business & Economic Statistics* **15**: 82–89.

Myers, S.C. 1984. Financial theory and financial strategy. *Interfaces* **14**: 126–137.

Nerkar, A. 2003. Old is gold? The value of temporal exploration in the creation of new knowledge. *Management Science* **49**: 211–229.

Pakes, A., Griliches, Z. 1984a. Patents and R&D at the firm level: A first look. In: Griliches, Z. (ed.). *R&D, Patents, and Productivity*. Chicago: University of Chicago Press.

Pakes, A., Griliches, Z. 1984b. Estimating distributed lags in short panels with an application to the specification of depreciation patterns capital stock constructs. *Review of Economic Studies* **51**: 243–262.

Patel, P., Pavitt, K., 1991. Large firms in the production of the world’s technology: An important case of non-globalization. *Journal of International Business Studies* **22**: 1–21.

Pisano, P.G. 1997. *The Development Factory*. *Unlocking the potential of process innovation*. Boston, MA: Harvard Business School Press.

Scotchmer, S. 1991. Standing on the shoulders of giants: Cumulative research and the patent law. *Journal of Economic Perspectives* **5**: 29–41.

Simons, J. 2006. Big Pharma’s new R&D center: The trash bin. *Fortune* **154** (12): 38.

Sørensen, J., Stuart, T. 2000. Aging, obsolescence, and organizational innovation. *Administrative Science Quarterly* **45**: 81–112.

Trajtenberg, M. 1990. A penny for your quotes: Patent citations and the value of innovations. *RAND Journal of Economics* **20**: 172–187.

Windmeijer, F. 2002. ExpEnd, a Gauss programme for non-linear GMM estimation of exponential models with endogenous regressors for cross section and panel data. Manuscript. Institute for Fiscal Studies.

Wooldridge, J.M. 1997. Multiplicative panel data models without the strict exogeneity assumption. *Econometric Theory* **13**: 667–679.

**Table 1 Descriptive Statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **S. D.** | **Min** | **Max** |
| R&D Expenditures\* | 504.75 | 587.83 | 1.51 | 2767.50 |
| Number of patents | 94.11 | 128.95 | 0 | 800 |
| First quartile number of patents | 14 |  |  |  |
| Median number of patents | 45 |  |  |  |
| Third quartile number of patents | 116 |  |  |  |
| Fraction with zero patents | 0.01 |  |  |  |
| Fraction with at least 100 patents | 0.29 |  |  |  |
| Pre-sample Mean | 69.30 | 91.76 | 0.10 | 362.80 |
| Equity\*\* | 4720.16 | 4825.73 | 7.08 | 17507.77 |
| European Firm | 0.30 | 0.46 | 0 | 1 |
| US Firm | 0.33 | 0.47 | 0 | 1 |
| Pharma Firm | 0.46 | 0.50 | 0 | 1 |

***Note***:

\* R&D Expenditures are inflation-adjusted in millions of year 2000 dollars.

\*\* We used the logarithm of book value of equity in year 2000 in millions of dollars, as a time-constant variable, to proxy for firm size. For those firms that fail to survive till year 2000, the end year value of equity is used and inflation-adjusted.

**Table 2 Correlation Matrix**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Patentst | Log R&Dt | Log R&Dt-1 | Log R&Dt-2 | Log R&Dt-3 | Log R&Dt-4 |
| Log R&Dt | 0.6101 |  |  |  |  |  |
| Log R&Dt-1 | 0.6167 | 0.9933 |  |  |  |  |
| Log R&Dt-2 | 0.6217 | 0.9839 | 0.9930 |  |  |  |
| Log R&Dt-3 | 0.6273 | 0.9749 | 0.9838 | 0.9933 |  |  |
| Log R&Dt-4 | 0.6299 | 0.9651 | 0.9749 | 0.9847 | 0.9934 |  |
| Log R&Dt-5 | 0.6318 | 0.9504 | 0.9624 | 0.9740 | 0.9826 | 0.9922 |

**Table 3 Estimates of the Knowledge Production Function from Patent Numbers**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Poisson****(random-effects)** | **Poisson****(fixed-effects)** | **Negative Binomial (random-effects)** | **Negative Binomial****(fixed-effects)** |
| Constant | -0.0403(0.5779) |  |  0.0048(0.3967) | -0.5668\*(0.2717) |
| Year Dummies | *Included* | *Included* | *Included* | *Included* |
| European Firm | -0.0164(0.2054) |  | -0.3471\*(0.1527) |  |
| US Firm |  0.7599\*\*\*(0.2000) |  |  0.1390(0.1584) |  |
| Pharma Firm | -0.6679\*\*\*(0.1700) |  | -0.5137\*\*\*(0.1187) |  |
| Log Equity | -0.1412**†**(0.0764) |  | -0.1523\*(0.0621) |  |
| Log R&D t |  0.3654\*\*\*(0.0362) |  0.3813\*\*\*(0.0363) |  0.2517\*\*(0.0943) |  0.2392\*(0.1028) |
| Log R&Dt-1 |  0.2493\*\*\*(0.0500) |  0.2449\*\*\*(0.0499) |  0.2937\*(0.1330) |  0.2432**†**(0.1415) |
| Log R&Dt-2 |  0.0365(0.0533) |  0.0445(0.0533) | -0.0164(0.1383) | -0.0416(0.1464) |
| Log R&Dt-3 |  0.1122\*(0.0536) |  0.1154\*(0.0536) |  0.0578(0.1379) | -0.0223(0.1439) |
| Log R&Dt-4 | -0.0522(0.0516) | -0.0493(0.0516) | -0.0948(0.1325) | -0.0854(0.1374) |
| Log R&Dt-5 |  0.2383\*\*\*(0.0372) |  0.2520\*\*\*(0.0373) |  0.2080\*(0.0934) |  0.1926\*(0.0956) |
| Sum of R&D Elasticity |   0.9495 |   0.9888 |  0.7000 |   0.5257 |
| Log Likelihood | -4498.0885 | -3888.3712 | -3327.3603 | -2713.6075 |
| Chi Square |  2339.20\*\*\* |  2212.43\*\*\* |  368.04\*\*\* |  226.49\*\*\* |

**Table 3 Continued**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **GMM** **І** | **PSM** **І** | **GMM** **II** |  **PSM** **II** |
| Constant |  |  3.7630\*\*\* (0.1001) |   |  1.7962\*\*\*(0.3536) |
| Year Dummies | *Included* | *Included* | *Included* |  |
| Log Pre-sample Mean |  |  0.6121\*\*\*(0.0660) |  |  0.3797\*\*(0.1383) |
| Patent t-1 |  |  |  0.2130\*\*\*(0.0172) |  0.9003\*\*\*(0.0439) |
| Log R&D t |  0.4443\*\*\*(0.0629) |  0.3058**†**(0.1583) |  0.7276\*\*\*(0.0315) |  0.4995\*\*(0.1529) |
| Log R&Dt-1 |  0.2664\*\*\*(0.0703) |  0.1646**†**(0.0981) |   |   |
| Log R&Dt-2 | -0.2992\*\*\*(0.0884) |  -0.0866(0.1393) |  |  |
| Log R&Dt-3 |  0.0248(0.0734) |  0.1394(0.1224) |   |  |
| Log R&Dt-4 |  0.1140(0.0699) | -0.1456(0.0940) |  |  |
| Log R&Dt-5 |  0.1698\*(0.0683) | -0.0707(0.1611) |   |   |
| Sum of R&D Elasticity |   0.7201 |   0.3069 |  |    |
| Sargan [p-value] |  55.5390[0.1139] |  |  69.0116[0.1991] |  |
| m1[p-value] | -2.5497[0.0108] |  | -4.0400[0.0001] |   |
| m2 [p-value] | -1.2863[0.1983] |  |  0.0049[0.9961] |  |

***Notes*:**

1. Standard errors are in parentheses. Year dummies are included but not shown, except for in PSM (II) estimation, for which we could not achieve convergence with the inclusion of fixed year effects.
2. **†** p<0.10, \* p<0.05, \*\* p<0.01, \*\*\* p<0.001.
3. *N* = 1,186 firm-year observations, the study has a slightly unbalanced panel sample of 85 firms covering the period from 1986 to 2000.
4. GMM (І) and PSM (І) provide estimates for Model (1), the multiplicative distributed lag model; GMM (II) and PSM (II) provide estimates for Model (2), the dynamic linear feedback model.
5. The diagnostic tests for GMM (I) and GMM (II) include: Sargan = the Sargan test for overidentifying restrictions, m1 = first-order serial correlation, m2 = second-order serial correlation; p-values are in brackets.

**APPENDIX A**

**Table A1 Autoregressive Estimates for Log R&Dt**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Equation** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** | **(12)** | **(13)** | **(14)** |
| Log R&Dt-1 |  0.993\*\*\*(0.005) |  1.158\*\*\*(0.044) |  1.144\*\*\*(0.049) |  1.152\*\*\*(0.049) |  1.175\*\*\*(0.051) |  1.154\*\*\*(0.044) |  1.147\*\*\*(0.045) |  1.130\*\*\*(0.046) |  1.138\*\*\*(0.048) |  1.170\*\*\*(0.049) |  1.154\*\*\*(0.045) |  1.140\*\*\*(0.047) |  1.151\*\*\*(0.048) |  1.180\*\*\*(0.049) |
| Log R&Dt-2 |  | -0.161\*\*\*(0.045) | -0.122(0.076) | -0.126(0.077) | -0.168\*(0.077) | -0.161\*\*\*(0.046) | -0.152\*\*(0.047) | -0.129\*\*(0.048) | -0.140\*\*(0.049) | -0.173\*\*(0.050) | -0.157\*\*(0.047) | -0.137\*\*(0.048) | -0.151\*\*(0.049) | -0.181\*\*\*(0.050) |
| Log R&Dt-3 |  |  | -0.022(0.041) |  0.061(0.071) |  0.029(0.076) |  |  |  |  |  |  |  |  |  |
| Log R&Dt-4 |  |  |  | -0.093(0.055) |  0.043(0.087) |  |  |  |  |  |  |  |  |  |
| Log R&Dt-5 |  |  |  |  | -0.081(0.056) |  |  |  |  |  |  |  |  |  |
| Log Patentt |  |  |  |  |  |  0.027\*(0.012) |  0.034\*(0.014) |  0.036\*(0.015) |  0.039\*\*(0.014) |  0.030\*\*(0.011) |  |  |  |  |
| Log Patentt-1 |  |  |  |  |  | -0.022(0.012) | -0.004(0.013) | -0.000(0.014) | -0.004(0.014) | -0.008(0.014) |  0.018(0.013) |  0.022(0.015) |  0.019(0.015) |  0.009(0.014) |
| Log Patentt-2 |  |  |  |  |  |  | -0.028(0.015) | -0.015(0.018) | -0.008(0.019) | -0.002(0.019) | -0.018(0.013) | -0.006(0.017) |  0.000(0.019) |  0.004(0.019) |
| Log Patentt-3 |  |  |  |  |  |  |  | -0.021(0.014) | -0.020(0.015) | -0.015(0.017) |  | -0.018(0.013) | -0.019(0.015) | -0.013(0.017) |
| Log Patentt-4 |  |  |  |  |  |  |  |  | -0.009(0.012) | -0.005(0.016) |  |  | -0.005(0.012) | -0.000(0.016) |
| Log Patentt-5 |  |  |  |  |  |  |  |  |  |  0.002(0.015) |  |  |  | -0.001(0.015) |
| R2 |  0.987 |  0.988 |  0.988 |  0.988 |  0.989 |  0.988 |  0.988 |  0.988 |  0.989 |  0.989 |  0.988 |  0.988 |  0.988 |  0.989 |

***Notes***:

1. All equations contain a separate intercept for each year.
2. Due to the presence of zero values, we added 1/3 to the patents variable before taking the logarithms.
3. Standard errors shown in parentheses are heteroskedastic-consistent estimates.
4. \* p<0.05, \*\* p<0.01, \*\*\* p<0.001.

**APPENDIX B**

1. **Multiplicative Distributed Lag Model [Model (1)]**

The conditional mean for a standard multiplicative distributed lag model for count panel data is of the form[[23]](#footnote-23):

 (B1)

where, and is a permanent scaling factor for the individual specific mean.

In the present study, the innovation model is written as,

 (B2)

where are the number of patents applied for and received by firm i in period t (i = 1, …, N; t = 1, …, T), denotes the natural logarithm of annual R&D expenditures, and is a vector of control variables.

 **GMM І**

In this study the generalized method of moment (GMM) estimator applies the Wooldridge’s (1997) quasi-differencing transformation[[24]](#footnote-24):

 (B3)

When are predetermined, the following moment conditions hold[[25]](#footnote-25),

. (B4)

 **PSM І**

The pre-sample mean (PSM) estimator solves the following moment conditions when estimating the multiplicative distributed lag model:

where is the pre-sample mean of y, TP is the number of pre-sample observations, and .

1. **Dynamic Linear Feedback Model[[26]](#footnote-26) [Model (2)]**

**Dynamic linear feedback model in the patents-R&D relationship:**

**An economic interpretation**

Instead of using Cobb-Douglas knowledge production function, R&D expenditures are combined through a specific parameterization of the CES knowledge production function to produce knowledge stock:

 (B6)

where is some latent measure of technological output of firm i in period t, k is a positive constant, R&D investment depreciates exponentially at rate , and captures the firm specific propensity to patents.

As observed patents are a noisy indicator of a firm’s technological output, we write

 (B7)

where are the number of patents applied for and received by firm i in period t, and .

Setting the returns to scale parameter equal to in (B6) and using (B7) results in,

. (B8)

Ignoring any feedback from patents to R&D expenditures, the long run steady state for firm i may be written as

 (B9)

thus may be interpreted as the long run elasticity of patents with respect to R&D, and is the short run elasticity.

Inverting (B8) results in,

 (B10)

 =

with .

From (B10), the conditional mean in the dynamic linear feedback model is defined as,

 (B11)

where

 are the number of patents applied for and received by firm i in period t (i = 1, …, N; t = 1, …, T), denotes the natural logarithm of annual R&D expenditures, and is a vector of control variables.

 **GMM II**

In this study the generalized method of moment (GMM) estimator applies the Wooldridge’s (1997) quasi-differencing transformation[[27]](#footnote-27):

When is predetermined, the following moment conditions hold

. (B13)

 **PSM II**

The pre-sample mean (PSM) estimator solves the following moment conditions when estimating the dynamic linear feedback model:

where is the pre-sample mean of y, TP is the number of pre-sample observations, and .

1. Internal R&D is usually measured as annual R&D expenditures by firms. A related question is to what extent patents serve as a good indicator of the output of R&D. Patents are directly related to inventiveness and represent an externally validated measure of technological novelty (Griliches, 1990). However, the use of patents as an economic indicator of knowledge increments has some limitations. For instance, not all inventions are patentable or patented, and the inventions that are patented differ greatly in their economic significance (Bound *et al.*, 1984; Griliches, 1990; Pakes and Griliches, 1984a). [↑](#footnote-ref-1)
2. The coefficient of the fifth year could be proxying for a series of small effects of the more basic research done six years ago or earlier, thus suggesting a lag “truncation” effect (Pakes and Griliches, 1984a). See Pakes and Griliches (1984b) for further discussion of this issue. [↑](#footnote-ref-2)
3. See Guo and Trivedi (2002) for a cross section analysis of the patents-R&D relationship. Their estimation results were in line with Hall *et al.* (1986). [↑](#footnote-ref-3)
4. Following Hall *et al.* (1986), a Granger causality test was performed in this study as well (see the results shown in Appendix A–Table A1). The current level of Log R&D was predicted with two lags of Log R&D (based on an approximate AR (2) specification) as well as contemporaneous and lagged Log Patents. As shown from column (6) through (10), the estimated coefficient on contemporaneous Log Patents was significant, but lagged Log Patents were of no help in predicting future R&D. The same behavioral pattern of lagged Log Patents was identified even when contemporaneous Log Patents was left out of the equation in columns (11)–(14). Thus, there was no evidence suggesting that past success in patenting led to an increase in a firm’s future R&D investments above and beyond that implied by its R&D history. [↑](#footnote-ref-4)
5. Chamberlain (1992) and Wooldridge (1997) developed a quasi-differenced GMM estimator that is consistent for count panel data models with predetermined regressors. This quasi-differenced GMM estimator has been applied to the analysis of the patents-R&D relationship by Montalvo (1997), Crépon and Duguet (1997), Cincera (1997), and Gurrmu and Pérez-Sebastián (2008). [↑](#footnote-ref-5)
6. A CBO Study: Research and Development in the Pharmaceutical Industry, Publication No. 2589, Congressional Budget Office, October 2006, available athttp://www.cbo.gov/ftpdocs/76xx/doc7615/10-02-DrugR-D.pdf. [↑](#footnote-ref-6)
7. Using Flemish R&D Survey data, Czarnitzki *et al.* (2009) provided empirical evidence on the differential contribution of research (‘R’) and development (‘D’) to patents and identified a high premium of research (‘R’), relative to overall R&D, towards firm patenting. [↑](#footnote-ref-7)
8. We are thankful to an anonymous referee’s valuable advice on this point. [↑](#footnote-ref-8)
9. Technical uncertainty refers to the likely costs and probabilities of accomplishing technical success. This form of uncertainty can only be reduced through investment by firms. Such technical uncertainty creates pressure on firms to invest immediately (Dixit and Pindyck, 1994; McGrath, 1997). [↑](#footnote-ref-9)
10. The sample for this study was drawn by referring to Compustat, Datastream, Amadeus, SIC reports, Ernst and Young’s Annual Biotech Industry Reports, Scrip’s Pharmaceutical Yearbook, amongst others. [↑](#footnote-ref-10)
11. For the sample firms in our study, total assets are a better indicator of scale change than capital stock, which was used in Hall *et al.* (1986). [↑](#footnote-ref-11)
12. See Guo and Trivedi (2002), Gurrmu and Pérez-Sebastián (2008) for the flexible techniques that are more desirable for highly overdispersed data. [↑](#footnote-ref-12)
13. We also used the logarithm of the annual values of equity to control for firm size. A one-year lag of this variable is included in all the model estimations. Since a few firm-year values of equity are negative, we changed the negative value into 0.000001 (million dollars) before taking the logarithms. The estimated results are basically similar. [↑](#footnote-ref-13)
14. The firm-specific effects $s\_{i} $are only included in the random-effects specifications. [↑](#footnote-ref-14)
15. There are reasons to believe that in many circumstances such correlations may exist. For instance, firms that have a higher propensity to patent for unobserved reasons may invest more in R&D as the return from R&D is higher than that from other investment projects (Montalvo, 1997). [↑](#footnote-ref-15)
16. The GMM estimation is inconsistent if the requirement on the instruments ( $z\_{it }$) is not satisfied (See Crépon and Duguet, 1997). Alternatively, $S\_{it}=(1,logRD\_{it-\left(τ+1\right)},logRD\_{it-\left(τ+2\right)},…,logRD\_{i1},w\_{t}^{'})$ can be used as the instruments, which do not require a restricted serial correlation of the residuals. The difference between these two sets of instruments is that the latter could lose efficiency in the presence of a restricted serial correlation of the residuals. We used $S\_{it}$ as well and the estimation results were basically similar. [↑](#footnote-ref-16)
17. Blundell *et al.* (2002) argue that inclusion of functions of the lagged dependent variable in the exponential function can lead to explosive series for patents or cause problems with transforming zero values. [↑](#footnote-ref-17)
18. Following Hausman *et al*. (1984) and Pakes and Griliches (1984a), the multiplicative distributed lag model [i.e., Model (1)] is constructed with five lagged R&D terms. [↑](#footnote-ref-18)
19. Our thanks go to Frank Windmeijer for his Gauss program, EXPEND, to estimate Model (1) and (2) using the GMM and PSM estimators (see Windmeijer, 2002). [↑](#footnote-ref-19)
20. In Montalvo (1997), the estimated coefficient on contemporaneous R&D was barely different from zero, whereas the first lag of R&D was positive and significant. Likewise, puzzling results were also found in the study by Gurrmu and Pérez-Sebastián (2008), where a negative and significant effect was identified at the fourth lag of R&D. [↑](#footnote-ref-20)
21. See Blundell *et al.* (2002), the PSM estimator performs better in Monte-Carlo simulations in comparison to the GMM estimator. However, there is an increase in the bias and root mean squared error for the PSM estimator of $β$ when the R&D series becomes very persistent. [↑](#footnote-ref-21)
22. We acknowledge that a firm’s knowledge production and innovation could arise from both its internal and external R&D investments. The focus in our paper is on the role of internal R&D as a basis for sustainable innovative capabilities of firms. The impacts of external R&D such as R&D alliances and acquisitions are beyond the scope of this study. As a robustness test, we also ran all the model estimations by including a firm’s external R&D investments, through R&D alliances and R&D acquisitions. R&D alliances are measured as the number of alliances between incumbent pharmaceutical firms and new biotechnology companies, in which a combined innovative activity or an exchange of technology is invloved. Likewise, R&D acquisitions are measured as the number of acquisitions undertaken by incumbent pharmaceuticals, which targetted new biotechnology companies for their R&D capabilities. A one-year lag was employed on both variables to alleviate a potential simultaneity bias. We found that the inclusion of external R&D does not alter the estimation results, including the lagged R&D effects and the U-shaped lag structure of the patents-R&D relationship, across the various estimations of model (1) and model (2). An exception is GMM І, in which the Sargan test rejects the null hypothesis and indicates that the model is misspecified. [↑](#footnote-ref-22)
23. The Cobb-Douglas knowledge production function is applied in this model. [↑](#footnote-ref-23)
24. An alternative is the Chamberlain (1992) quasi-differencing transformation:

$s\_{it}=y\_{it}\frac{μ\_{it-1}}{μ\_{it}}-y\_{it-1}$. [↑](#footnote-ref-24)
25. $z\_{it}=(1,x\_{it-\left(s+1\right)},x\_{it-\left(s+2\right)},…,x\_{i1},control^{'})$ is used as the instruments, where s ( s = 0, …,$ τ$ ) represents the extent to which the $z\_{it}$ are weakly exogenous, i.e., s = 0 (s = 1, …,$ τ$) means that$ x\_{it-1} $($x\_{it-2}$ , …,$ x\_{it-\left(τ+1\right)}$) in the$ z\_{it}$ (Cincera, 1997). [↑](#footnote-ref-25)
26. See Blundell *et al.* (2002). [↑](#footnote-ref-26)
27. An alternative is the Chamberlain’s (1992) quasi-differencing transformation as follows:

$s\_{it}=(y\_{it}-γy\_{it-1})\frac{μ\_{it-1}}{μ\_{it}}-(y\_{it-1}-γy\_{it-2})$. [↑](#footnote-ref-27)