

Corporate Science, Innovation, and Firm Value

Markus Simeth ^{a, b}

Michele Cincera ^c

This version: May 2015

Paper forthcoming in Management Science

Abstract

Many firms actively disclose research findings in scientific peer-reviewed journals. The literature highlights several potential benefits of such scientific boundary-spanning activities, including privileged access to academic information networks. However, scientific disclosure may lead to unintended knowledge spillovers. It remains unclear whether active engagement in science leads to higher returns. This paper investigates the impact of scientific activities on the firm's market value, using accounting data for US firms and matched patent and scientific publication data. We find evidence for the positive impact of scientific publications on a firm's market value beyond the effects of R&D, patent stocks and patent quality and also document heterogeneity with respect to this impact between different industrial sectors.

Keywords: R&D, Industrial science, Scientific disclosure, Tobin's Q, Knowledge disclosure, Innovation

^a KU Leuven, Department of Managerial Economics, Strategy & Innovation, Belgium

^b Universidad Carlos III de Madrid, Department of Business Administration, Spain.

^c Université Libre de Bruxelles (ULB), Solvay Brussels School of Economics and Management, iCite, Belgium

Corresponding author: Markus Simeth, KU Leuven, Department of Managerial Economics, Strategy and Innovation, Naamsestraat 69, B-3000 Leuven. Email: markus.simeth@kuleuven.be

1 Introduction

Research and Development (R&D) is an important source of competitive advantage for companies. In particular, investments in upstream research lead to the creation of new knowledge inside a firm and also facilitate the absorption of academic institutions' and researchers' latest knowledge (Cohen and Levinthal 1989; Rosenberg 1990). Consequently, firms investing in research may produce radical inventions which promise greater financial returns. Interestingly, many companies involved in R&D not only absorb external academic knowledge but actively contribute to open science by disclosing their own findings in scientific peer-reviewed journals (Hicks 1995).

The literature highlights several mechanisms through which firms who adopt academic disclosure practices might benefit. Broadly summarizing, advantages include increased capacity to absorb external knowledge, the realization of signaling benefits which facilitate access to valuable research inputs such as hiring of PhD graduates, and the promotion of science-based products among professional customers (Hicks 1995; Stern 2004; Polidoro and Theeke 2012). However, the dissemination of research findings in scientific journals may also lead to unintended knowledge spillovers that facilitate imitation by competitors and other specific costs associated with the disclosure process (Arrow 1962; Cockburn et al. 1999; Kinney et al. 2004; Stuart and Liu 2014). Although empirical studies document that firms with scientific publications lodge higher quality patents (Cockburn and Henderson 1998; Gittelman and Kogut 2003; Fabrizio 2009), these studies disregard the potential spillover costs caused by scientific disclosure. It remains unclear therefore whether increased R&D productivity by a firm genuinely translates into higher financial profitability.

In this paper, we address this question and investigate the profitability implications of scientific disclosure. In order to capture all positive and negative effects, we apply a market value approach (Griliches 1981; Jaffe 1986; Hall 1993; Hall et al. 2005). Financial markets react immediately to new information and anticipate future commercial implications. Thus, this approach avoids making assumptions about time lags between knowledge creation, disclosure effects, and the realization of financial returns. We separate - conceptually and empirically - the creation of scientific outcomes from the voluntary disclosure of these outcomes in scientific peer-reviewed journals. We also provide initial insights regarding the broad mechanisms through which scientific publications may create value. A further distinctive contribution is that unlike the majority of studies dealing with science in the corporate context which refer to the biotechnology and pharmaceutical industries, the present paper considers all high-technology sectors. Since sectors differ according to their appropriability conditions (Levin et al. 1987; Cohen et al. 2000) which may moderate spillover risks of disclosure, we examine sector heterogeneity in greater detail.

Our empirical analysis uses firm-level information from the US *Compustat* database, and matched scientific publication and patent data for a representative sample of 1,739 stock-market listed firms

(9,920 firm-year observations) from all high-technology sectors, covering the period 1996-2006. The results of our econometric analysis suggest a positive relationship between the scientific publications of firms and their Tobin's Q. We find evidence that disclosure in high-quality journals, which indicates active involvement in science, is especially valuable in terms of higher stock market premium. However, we detect a certain degree of heterogeneity across high-technology sectors: whereas scientific publications are associated with positive returns in the instruments sectors, and to some extent also in pharmaceutical and biotechnology, we find a negative but non-significant tendency in the ICT sectors.

2 Theoretical background

This paper builds on the already large body of literature examining the relationship between R&D and market value, and on recent work investigating the scientific activities of profit-oriented firms. With regard to the former, several studies have examined the impact of R&D expenditures, patent stocks and patent quality indicators on firm value (see Hall 2000; Czarnitzki et al. 2006, for surveys; Hall et al. 2005). The majority highlights a positive effect of R&D and patent stocks on the firms' market value, including a premium for high-quality patents. More recent studies have taken into account the strength of the appropriability regime, complementary assets and knowledge spillovers as moderating factors on this impact (Cockburn and Griliches 1988; MacGahan and Silverman 2006; Belenzon 2012). To date however, few studies have considered heterogeneity of R&D activities across firms. One exception to this in accounting literature is the study by Deng et al. (1999) which found a positive effect of executing scientific research on firm value. However that study did not analyze voluntary contributions of firms to the stock of open science. Instead, the authors relied on non-patent literature references in a firm's list of patents as a proxy for scientific research. We build on their work and offer a more complete perspective, with a special focus on the impact of the active participation of firms in the scientific community, tangibly reflected in scientific publications.

From a conceptual point of view, there are several reasons why the scientific activities of firms may have implications for firm profitability. Firms engaging in (basic) research may develop superior capabilities, in that they may be able to combine technologically distant knowledge to create inventions which are more valuable than those created by other firms. Such firms are potentially able to absorb external knowledge and identify promising trajectories for applied research and experimental development, leading to superior inventive outcomes (Rosenberg 1990; Cohen and Levinthal 1989; Gittelman and Kogut 2003; Fleming and Sorenson 2004; Fabrizio 2009). While these studies also suggest a positive relationship between research investments and firm profitability, the potential costs of basic research are not considered. In particular, basic research findings do not necessarily lead to high commercial success since translating findings into concrete products can be challenging (see Nelson 1959; Rosenberg 1990; Cockburn et al. 1999; Pisano 2006).

Beyond the creation of scientific knowledge, many firms also voluntarily disclose their research results in scientific journals. Such involvement in open science may create value through several mechanisms associated with access to upstream knowledge and appropriation (see Penin 2007 for a review). With regard to upstream knowledge, disclosing research results may not only increase the absorptive capacity (Cockburn and Henderson 1998), but may facilitate knowledge sourcing from academic scientists, thanks to reciprocity in knowledge exchange. More specifically, academic scientists may be reluctant to interact with scientists from firms unless the latter share valuable research results (Hicks 1995; Simeth and Raffo 2013). Moreover, scientific publications can be leveraged as a human resource instrument with regard to PhD graduates. Many scientists who consider the private sector as a career option value the possibility of continuing publication activities, and may seek out firms engaging in this practice (Stern 2004; Sauermann and Roach 2013). Scientific publications are also considered a potential appropriability device. Disclosure may encourage the adoption of science-based products such as medical drugs or instruments (Azoulay 2002; Polidoro and Theeke 2012). Clinicians, as users of medical instruments, influence procurement decisions in university hospitals and are keen to understand about the technological characteristics of the products. Manufacturers of such devices might gain greater credibility among these clinicians by disclosing related information in scientific peer-reviewed journals. Another appropriation-related use of scientific documents is the establishment of prior art through disclosure, the aim being to hamper patenting strategies of competitors working on similar inventions (De Fraja 1993; Parchomovsky 2000; Della Malva and Hussinger 2012).

However, the publication process has related costs. The most obvious of these are unintended knowledge spillovers that may enable competing firms to learn from the disclosed knowledge (Arrow 1962). Knowledge spillovers may reduce the cost of imitation for competitors, offer the latter insights into future trends and facilitate their exploration of alternative technological trajectories. In addition, the publication process itself can lead to opportunity costs since a firm's researchers have to prepare documents which meet the publishing requirements set down by journals, interact with referees and codify tacit knowledge (Kinney et al. 2004). Furthermore, firm scientists who publish are more visible to competing firms, which may impose the need for retention strategies as their external employment options increase (Stuart and Liu 2014; Kim and Marschke 2005). In summary, the positive and negative effects of the relationship between corporate science and firm value underline the need for further study in this area.

3 Data and Methodology

3.1 Data sources

Our analysis was based on a sample of US public companies from high-technology industries, as defined by the OECD (OECD 2011), including pharmaceuticals and biotechnology, telecommunication

equipment and semiconductors, aircraft, as well as scientific and medical instruments. We also included the chemicals sector since it is often recognized as science-oriented. These industries were a natural choice as they draw on scientific knowledge as an input factor, while their publication output is varied (Cohen et al. 2002; Simeth and Raffo 2013). Firm-level data came from Standard and Poor's *Compustat*, patent data from the *Worldwide Patent Statistical Database (Patstat)*, version April 2014) provided by the European Patent office, while publication data came from Elsevier's *Scopus* database. As coverage of *Scopus* changed considerably in 1996, we were obliged not to include earlier periods in our analysis.¹ Moreover, the patent information concerning the United States Patent and Trademark Office (USPTO) patents contains the caveat that applications filed before 29th November 2000 but not granted are unobservable.² In order to keep the numbers of patents comparable over time, we therefore only considered patent applications where we observed the grant within seven years after the application date. This allowed for the inclusion of patent data until 2006. Consequently, the sample covers the time period 1996-2006, and firms were considered for selection based upon the criterion that they invested in R&D in at least one of the years during this period.

Combining firm-level data with publication and patent information may lead to errors arising from name-based matching procedures. To achieve high recall rates of publication and patent numbers, we carefully pre-cleaned all firm names by correcting misspellings and removing abbreviations denoting corporate structure (e.g. "Inc."). The publication and patent matching processes differed as scientific publications were retrieved manually from the *Scopus* online database, whereas patent data was matched using an offline source. The former is a more time-consuming process than "automatic" offline matching, and the latter requires extensive manual cleaning of the algorithm-based matching results. Manual retrieval of publication data enabled us to directly inspect hits and immediately exclude other institutions with an identical name. The process is explained in further detail in Appendix A.1. For the offline matching procedure with patent data from the *Patstat* database we first pre-tested all names to detect problematic cases. Firms with ambiguous names were either excluded directly or marked for detailed manual inspection once the actual matching took place. Second, we applied name standardization routines both to the names of our sample firms and the applicant field in the *Patstat* database. Third, after testing several matching algorithms, we identified the one which achieved the highest recall rate while simultaneously limiting false positive hits (see Raffo and Lhuillery 2009). The resulting matches were checked manually with particular focus on firms with atypical input-output ratios and those identified in the pre-tests as being uncertain. In order to cross-check the quality of our patent-

¹ *Scopus* extended its journal coverage considerably in 1996, potentially inducing bias to the econometric estimations. See http://www.info.sciverse.com/UserFiles/2508.SciVerse.Scopus_Facts_Figures%28LR%29.pdf.

² The policy change was introduced by the American Inventors Protection Act (AIPA) and concerns all patents that seek protection in the United States (except patents that enter into the PCT procedure for worldwide protection and "opt-out" cases). As of 29th November 2000 all patent applications are disclosed with the search report of the patent examiner after 18 months.

matching process, we also combined our data with the Compustat-NBER dataset (see Hall et al. 2001), performing regressions on the patent numbers based first on our matching and then on Compustat-NBER matching. The results were consistent. Since the Compustat-NBER dataset also considers static ownership structure (for the year 1989), this exercise suggests that our results were not affected by our inability to take into account subsidiaries of the studied companies in the present study. However, it is still theoretically possible that we missed a relevant number of publications when not considering subsidiaries. To mitigate such concerns, we searched for information on subsidiaries for some of the investigated sectors (Biotech SIC 2845 & 2846, Semiconductors SIC 3674, Instruments SIC 3841 & 3842) and downloaded publications taking into account the names of the subsidiaries using the SEC security filings (report “10-K”). The corresponding regressions suggested that no bias arose from the non-consideration of subsidiaries in our analysis (see Appendix A.2).

To reduce biases originating from Merger & Acquisition (M&A) activities, measurement errors and “atypical” firms (e.g. specialized R&D firms with a majority of shares owned by business groups), several filters were applied to the sample. First, to limit potential biases from M&A events, firms with large changes in book value were identified based on the criteria of an increase of more than 300% or decrease of more than 75% over a period of two consecutive years (Griliches 1981; Hall and Oriani 2006; Aldieri and Cincera 2009). We only dropped the firm-year observation when a large change occurred. In such cases, we treated the firm in subsequent years as a new firm (Griliches and Mairesse 1984). Excluding firms involved in M&A activities entirely could have led to selection bias since M&A activities are presumably often based on successful R&D operations. To address measurement errors and inaccurate initial knowledge stock computations, we also dropped observations with extremely high (the top 1%) knowledge-to-asset ratios.³ Finally, we excluded firms with fewer than 10 employees, imposed a minimum sales amount of 500,000 USD, and removed firms with an R&D/Sales ratio higher than one. A firm with no sales generates a high degree of uncertainty among investors in terms of its survival prospects. Such firms are therefore at a different structural stage than those already generating a cash flow, the result likely being heterogeneous treatment on the stock market. In our study, since the R&D/Sales ratio filter affected a notable number of observations, we removed it as a robustness test (see Section 4.4). The final sample size contained 9,920 firm-year observations, composed as follows:

-- Insert Table 1 here --

Finally, all financial amounts were adjusted for inflation using the Gross Domestic Product (GDP) deflator.

³ The three measures concerned are *R&D/A*, *PAT/R&D*, *PUB/R&D* which are explained in detail in section 3.2.

3.2 Variables

To construct our dependent variable, we applied the methodology followed by previous related research and used the Tobin's Q which represents the ratio of a firm's market value to its book value. Market value is the addition of the market value of the equity and the market value of the debts. The former is calculated by the number of outstanding shares multiplied by the stock price at the end of the fiscal year, whereas the latter is approximated using the book value of liabilities (Blundell et al. 1999; Hall and Oriani 2006, Ceccagnoli 2009). The firm's book value is represented by its assets at the end of the fiscal year. With respect to the independent variables, one core measure is the firm's R&D investments which reflect overall commitment to knowledge production. The investments in R&D take place in a particular period, but the returns on these investments may last much longer. Therefore, we introduced a stock measure of R&D. Since knowledge becomes obsolete because of ongoing technological development, we applied the frequently used perpetual inventory method with an annual 15% depreciation (δ) rate (Griliches and Mairesse 1984; Hall et al. 2005). In the absence of a strong theoretical justification for assuming different rates for the other knowledge-related measures, we not only computed the R&D stock but also the publication and patent stock indicators with the 15% rate as follows:

$$R\&D\ STOCK_t = R\&D_t + (1 - \delta)R\&D\ STOCK_{t-1} \quad (1)$$

Although the computation of the R&D stocks is technically straightforward, assumptions have to be made regarding the initial stock ($R\&D\ STOCK_{t_0}$) which remains partly unobserved. In our study, we applied a standardized growth rate (g) for R&D and the other knowledge stock measures of 8%, given that our sample contained only high-technology firms (Hall and Oriani 2006; Hall et al. 2007).⁴ All knowledge-related measures discussed below are accordingly constructed as stocks. We constructed three measures directly related to the firms' scientific operations. First, we use a patent-based indicator to measure outcomes from basic research, and thus indirectly also measure a firm's potential to create scientific publications (*SCIPAT*). It is assumed that the existence of a scientific document in the backward references of a patent indicates a science-based patent (Narin et al. 1997; Deng et al. 1999). We identified scientific documents in the Non-Patent-Literature (NPL) section of patents' backward references using specific keywords and character combinations collected via extensive screening of the raw NPL information. Since this measure captured research outcomes but did not depend on the observation of scientific contributions to academic journals, the use of *SCIPAT*, in combination with the publication-based indicators, enabled us to differentiate between research outcomes and disclosure

⁴ This applies specifically to firms that do not have a long pre-sample record in *Compustat*. The stocks can be (at least partly) computed with observed values for firms with IPO's before 1996. In formal terms, the initial R&D stock is approximated as follows: $R\&D\ STOCK_{t_0} = R\&D_{t_0} / (\delta + g)$

effects.⁵ Although this measure of basic research relies on patented outcomes, we nonetheless considered it to be a reasonable measure, given that our sample comprised only firms from high-tech sectors with a high propensity to patent.

Second, we captured the number of voluntary scientific contributions by computing a stock measure based on the number of scientific papers published by each firm (*PUB*). Given our use of *SCIPAT* to measure basic research outcomes, the publication stock should not only reflect research productivity but also effects deriving from the disclosure of results. Moreover, to consider the heterogeneity of scientific contributions, we introduced a second measure reflecting their academic quality (*TOPPUB*). If firms are able to publish in prestigious journals, it is likely that the degree of involvement in the scientific community will be higher. Publishing in highly ranked journals is also less likely to reflect appropriation motives since the publication process is riskier with higher likelihoods of rejections and delays imposed by referees' requests for further experiments before accepting the submitted article. In order to overcome potentially imperfect corrections for field effects and inaccurate weightings of articles by impact factors, we decided to rely on a simpler dichotomous distinction by identifying the top 10% of journals based on the impact factor distribution within the five meta-disciplines: Life Sciences, Physical Sciences and Engineering, Social Sciences, Health Sciences and General/Interdisciplinary. The computed stock variable therefore only considered scientific documents published in these top journals (see Appendix A.3).

In line with previous literature, we included two measures reflecting the amount of inventive outputs and their quality. The absolute amount of inventive outcomes is represented by patent stocks (*PAT*) and the quality of the inventive outcomes (*FWDCIT*) is measured using forward citation counts (Hall et al. 2005). To deal with the truncation problem of citation counts, we considered only those citations that occurred within a five-year window after the priority date of our focal patents (Lanjouw and Schankerman 2004; Marco 2007). Finally, we included sector and year dummy variables to take into account heterogeneous market valuations across industries and time (Cockburn and Griliches 1988; Hall et al. 2005).

3.3 Model and estimation techniques

In this paper, we analyze the relative market value of firms (Tobin's Q) as a function of their knowledge stocks. We rely on the well-established market value function (Griliches 1981, Hall et al. 2005) which

⁵ Taking into account the possibility that *SCIPAT* might not exhaustively capture the basic research outcomes, we also included a stock of "original" patents based on the patent originality measure proposed by Trajtenberg et al. (1997). The results of this specification are reported in the appendix A.4 and are consistent with those of the main model.

regards tangible (A_{it}) and intangible assets (K_{it}) as additives (“hedonic model”). The function can be formalized as follows:

$$V_{it}(A_{it}, K_{it}) = q_{it} (A_{it}, \gamma K_{it})^\sigma \quad (2)$$

In this equation, q_{it} represents the valuation coefficient of a firm’s assets, and the parameter γ allows for the eventuality that knowledge assets are valued differently from physical assets. The valuation coefficient q_{it} may vary across time and industries and also contains a firm-specific component. The factor σ represents scale effects and is assumed to equal 1, which is also confirmed in our data (Hall et al. 2005). Applying logarithms on both sides and moving tangible assets to the left-hand side of the equation yields the following expression:

$$\text{Log} \left(\frac{V_{it}}{A_{it}} \right) = \log Q = \log q_{it} + \log \left(1 + \gamma \frac{K_{it}}{A_{it}} \right) + e_{it} \quad (3)$$

Following our theoretical discussion, we separate the knowledge intangibles of firms into R&D, patent stocks, and publication stocks. Since patents and scientific publications can be regarded as direct outcomes of R&D inputs, we introduce these additional measures as ratios denominated by R&D expenditures (Hall et al. 2005; Hall and Oriani 2006; Hall et al. 2007), formalized in Equation (4):

$$\text{Log} \left(\frac{V_{it}}{A_{it}} \right) = \log Q = \log q_{it} + \log \left(1 + \gamma_1 \frac{R\&D_{it}}{A_{it}} + \gamma_2 \frac{PAT_{it}}{R\&D_{it}} + \gamma_3 \frac{PUB_{it}}{R\&D_{it}} \right) + e_{it} \quad (4)$$

In our empirical setting, $R\&D$, PAT and PUB represent the respective stock measures computed using the perpetual inventory method with an assumed depreciation rate of 15%. In the full specification, we distinguish between the impacts of (*basic*) *research* and *disclosure* effects using further measures that are introduced as ratios and which are orthogonal to the main variables. These further variables reflect science-based inventions ($SCIPAT/PAT$), the quality of the inventive output ($FWDCIT/PAT$), and the academic impact of scientific output ($TOPPUB/PUB$). Consequently, the extended model of Equation (4) can be written as:

$$\begin{aligned} \text{Log} \left(\frac{V_{it}}{A_{it}} \right) = \log Q = \log q_{it} + \log \left(1 + \gamma_1 \frac{R\&D_{it}}{A_{it}} + \gamma_2 \frac{PAT_{it}}{R\&D_{it}} + \gamma_3 \frac{PUB_{it}}{R\&D_{it}} + \gamma_4 \frac{SCIPAT_{it}}{PAT_{it}} + \right. \\ \left. + \gamma_5 \frac{FWDCIT_{it}}{PAT_{it}} + \gamma_6 \frac{TOPPUB_{it}}{PUB_{it}} \right) + e_{it} \end{aligned} \quad (5)$$

Equations (4) and (5) can be directly estimated using Non-Linear Least Squares (NLLS). Instead, OLS regression models can be applied only by using the approximation $\log(1+x) \sim x$. Nevertheless, higher values of x lead to imprecisions in the estimated outcome. An important concern is the potential presence of unobserved firm heterogeneity. However, independent variables are potentially predetermined since past firm valuations may influence current investment decisions related to R&D strategies. This eventuality violates the strict exogeneity assumption for the “within” estimator. Moreover, unlike the rather volatile Tobin’s Q indicator, the innovation-related variables change slowly over time and are thus highly correlated with the firm-specific effect. Consequently, a “within” fixed-effects estimator may rather exacerbate estimation problems than allow valid estimates to be obtained (see also Hall et al. 2005). A potential solution, suggested by Blundell et al. (1999), is the introduction of a pre-stock average of the firm’s market value as an additional regressor. We tested this approach, but it led to a substantially decreased sample since many firms do not have a sufficiently long stock market history. Consequently, like most of the related empirical literature, we decided to focus on NLLS regression models.

3.4 Descriptive statistics

In Table 2 we provide an overview of the mean and median values as well as the standard deviation of the regression variables and selected additional measures that provide information about firm characteristics.

-- Insert Table 2 here --

Since all firms were stock market-listed, the sample consisted predominantly of medium- and large-sized firms. However, median values suggest considerable heterogeneity among our sample of firms. The median values for annual R&D expenditures were 8.36 million USD, 79.41 million USD for sales and 340 for the number of employees, whereas the mean values were much higher, indicating the presence of both large but also medium-sized and small firms. The average Tobin’s Q ratio was quite high at 2.65. In 12.3% of the firm-year observations, the market valuation was below the book value. With regard to trends in publication and patent outputs, the share of both publishing and patenting firms increased over time. Time trends are shown in Figures 1 and 2.

-- Insert Figure 1 & 2 here --

The average number of publications and patents by a firm both increased over time. At the beginning of the sample period the number of patents continuously increased while the number of

publications remained stagnant, before starting to increase in 2002. Interestingly, the average level of publication and patent activity did not differ greatly, which is remarkable given that firms are, by definition, not concerned with contributions to the stock of scientific knowledge *per se*. With regard to the share of firms with at least one publication or patent in a given year, we obtained the following picture as reported in Figure 2: the share of publishing firms increases from 37% in 1996 to 53% in 2006, whereas the share of patenting firms increased from 46% in 1996 to 60% in 2006. Output types appeared to be complementary since the majority of publishing firms in a given year also filed a patent application. Over time, there was an increase in the strategy to patent and publish simultaneously. In 2006 however, 14% of all firms only published, 17% only created patents, and 26% of firms neither created patents nor published.

4 Econometric results and discussion

4.1 Full sample estimations

The core specifications of the econometric analysis are shown in Table 3. Considering the elements highlighted in the previous section, we estimated Tobin's Q with both non-linear and linear regression models, but ultimately focused on the interpretation of the former since non-linear models enable the theoretical model to be directly represented. In order to evaluate the quantitative impact of our variables in the NLLS regressions, we also computed semi-elasticities (at the mean values of the variables) which are reported in complementary columns for the baseline (2) and full (6) NLLS regression models.⁶

-- Insert Table 3 here --

In columns (1)-(2), we estimated baseline specifications using NLLS and OLS estimators. All three knowledge measures representing R&D ($R\&D/A$), patent ($PAT/R\&D$) and publication ($PUB/R\&D$) stocks had a positive and significant effect on Tobin's Q, the latter two having similar magnitudes. One additional patent per million dollars of R&D was associated with an increase of 9% in Tobin's Q, and 10% for an additional publication per million dollars of R&D. This result provided a first indication that scientific activities have an important impact on the market value of firms. In order to obtain a more detailed picture, we introduced the additional measures described in Section 3.3. In columns (3)-(5), we first separately added the three variables $FWDCIT/PAT$, $SCIPAT/PAT$ and $TOPPUB/PUB$ to the baseline model. All three measures exhibited a value premium for Tobin's Q. The main variables remained statistically significant when the supplementary patent-based variables were included.

⁶ Semi-elasticities can be obtained by calculating the derivative of the estimated market value equation with regard to the variable of interest. See Hall et al. (2005) for further details.

However, when adding the measure capturing the share of publications in prestigious journals (*TOPPUB/PUB*) in column (5), the magnitude of the publication stock *PUB/R&D* decreased, and the variable only remaining significant at the 10% level. In columns (6)-(7), we report the regression results with the full variable set. In the NLLS specification (6), it can be seen that R&D stocks, patent stocks and patent citation stocks (*FWDCIT/PAT*) had a positive impact, which is in line with findings from previous studies (e.g. Hall et al. 2005). While publication stocks (*PUB/R&D*) were only significant at the 10% level, we detected a strong positive and highly significant effect of publications in top journals. The increase of the variable by one unit – which is equivalent to a change from none of a firm’s publications being in top journals to all of the firm’s publications being in top journals - increased Tobin’s Q by approximately 12%.⁷ Interestingly, the impact of science-based patents (*SCIPAT/PAT*) became substantially weaker in comparison to column (4), with the positive effect of a one unit change decreasing from 11% to 4%, and the variable becoming statistically insignificant once the share of publications in top journals was taken into account.

Based on these regression results, the question arises as to whether the positive effects of scientific activities which we found reflect (i) the successful creation of science-based knowledge outcomes or benefits that derive from the disclosure of these outcomes, (ii) signaling benefits to academic audiences or the use of publications as a device to support appropriation. We would like to stress that we cannot provide conclusive evidence concerning these mechanisms with our measures but can only offer some indications which must be treated with caution. In the full model presented in column (6), positive effects for scientific publication stocks and especially top journal publications were observed despite the presence of science-based patents. While being highly significant and exhibiting a notable magnitude in column (4), the measure of science-based patents was no longer significant in the full specification. This indicates that the use and absorption of basic research has a positive impact on a firm’s market value. However, active involvement in science which is especially represented by publications in top journals, generates a considerable market value premium beyond the successful absorption of scientific knowledge. The relevance of publication quality also has implications with regard to point (ii) above and the underlying value-creating mechanisms of scientific disclosure. Firms that disclose knowledge for defensive purposes have an interest in keeping control over the timing of disclosure. As a consequence, there are incentives for firms to target lower impact journals as, presumably, the likelihood of acceptance in high-impact journals is lower and the time taken to publish longer. Therefore, if this mechanism is value-enhancing, we would expect a stronger effect of the main publication stock *PUB/R&D* in the full model. However, the particularly strong effect of publications in high-impact journals points more to value deriving from facilitated interactions with academic partners. In other

⁷ Caution must be exercised when comparing the magnitudes between *PUB/R&D* and *TOPPUB/PUB* due to the different units of the variables. Therefore we also calculated the impact of a change by one standard deviation. The result of this exercise also suggested that publications in prestigious journals are associated with a particularly strong increase in Tobin’s Q (3.7% vs. 1.6%).

words, firms actively participating in the scientific community seem to have easier access to valuable knowledge and to highly qualified graduates, which results in a higher Tobin's Q.

4.2 Sector analysis

Since the association between scientific activities and firm value may differ across industries, especially in the light of the relative risks of knowledge spillovers, we analyzed the individual sectors in greater detail. In Table 4, we report the results of subsample regressions with the meta-sectors Biotechnology and Pharmaceuticals (hereafter BIO-PHARM), Information & Communication technologies (ICT), Scientific and Medical instruments (INST), and Chemicals (CHEM).

-- Insert Table 4 here --

Starting with the baseline specification, it can be seen that the effect of scientific publication stocks differs across the meta-sectors. Although we detected a positive but non-significant sign in BIO-PHARM and CHEM, a negative and non-significant effect was found in ICT, while a positive and significant effect was observed for the INST sectors. Estimating the regression model with the full set of variables revealed interesting patterns. In BIO-PHARM, the publication quality indicator was significant and of quite a strong magnitude, whereas the variables reflecting the main publication stock and science-based patents did not have such an impact. This result suggests that as long as a firm is not seriously committing to Open Science, scientific activities will not be sufficient to generate a value premium. We observed a similar result for the INST subsample, where the *TOPPUB/PUB* measure had a strong positive effect too. However, the main publication variable *PUB/R&D* also showed a rather strong magnitude in comparison to the other subsamples, and remained significant. This finding suggests that in addition to upstream signaling, publications may create value through mechanisms other than establishing knowledge flows with the scientific community. One plausible interpretation for this finding lies in marketing effects, since firms in instrument sectors sell to scientifically-trained professional customers. In order to establish credibility among this audience with regard to the latest science-based products, publishing in high-impact journals is presumably less crucial. In the ICT domain, we saw a different picture. Publication stocks were negative and the share of top impact journal publications also had a negative sign. Interestingly, in addition to the positive effects of patent stocks and patent quality, the share of science-based patents was positive and significant. In other words, absorbing scientific knowledge and achieving science-based inventions may create value, whereas active dissemination of results and participation in the scientific community do not. For the CHEM sector, we only observed a positive effect of R&D on the market value. However, there was no premium for patent and publication stocks or for science-based patents.

These obtained differences across sectors indicate that appropriability conditions have a notable influence on the returns of scientific disclosure strategies. In sectors with strong legal appropriation, like the medical instrument and biopharmaceutical industries, we observed positive net returns from scientific publications, whereas in the ICT domain, which can be characterized as having a weaker appropriability regime due to its complex and cumulative nature, no premium was observed (see Levin et al. 1987; Cohen et al. 2000; Hall and Ziedonis 2001). This finding indicates that in the ICT sectors, scientific disclosure leads to knowledge spillovers which outweigh any benefits. The result for the Chemicals sector, where appropriability conditions are stronger than those in the ICT sector, can be explained by the fact that in the former, academic knowledge is only of moderate importance as an input. Therefore, the potential benefits of scientific disclosure in terms of signaling might be limited in the first place.

4.3 Robustness tests

Beyond considering subsidiary publications for a subsample of firms, as reported in Appendix A.2, we performed further robustness checks, the most important of which are displayed in Appendix A.4. These robustness tests did not change our main models' results. Consequently, we will only describe them very briefly. We list several tests and for ease of reading we shall assign them numbers that correspond to the respective columns in Appendix A.4. (1) We ran fixed-effect regression models. As discussed in Section 3.3, such models must be interpreted with caution in our context. However, the results suggested that scientific activities do indeed lead to higher firm values. (2) We removed the filter which initially excluded firms whose R&D/Sales ratio was larger than one since this filter was perhaps overly restrictive. However, our results after removing this filter did not change. (3) In order to control for unobserved selection mechanisms for publication in terms of local spillovers and competition, we included state dummies. (4)-(5) Since scientific backward references may not exhaustively capture basic research outcomes in firms, we included the patent originality variable as suggested by Trajtenberg et al. (1997). This variable behaves very similarly to science-based patents, and supports the interpretation that the results for publication stocks do not only reflect research productivity effects but also benefits deriving from the active involvement of firms in science. (6)-(8) We introduced additional control variables, namely the amount of sales and the growth of sales, sector-level patent and publication propensities, and the share of firm's profit on the industry profit. These controls should capture the effect of unobserved intangibles potentially correlated with our knowledge intangibles, of the abilities of firms to appropriate returns from R&D, and in the case of the sector-level variables, of incentives to engage in publication. Our results did not change when these additional controls were included. (9) We used the patent counts from the Compustat-NBER matching (Hall et al. 2001; Cockburn et al. 2009) to verify the quality of our own matching. The regression outputs were very similar. (10)-(11) We modelled the

market value function in a different way by including the publication stock to patent stock ratio (i.e. instead of publication stock to R&D stock ratio) for the subsample of patenting firms. Such a specification might be an adequate measure of openness if publications and patents reflect the same knowledge. The variable PUB/PAT was found to be positive and significant, supporting the interpretation of the core models that publication provides additional value beyond patented knowledge. (12) We included an additional measure of the share of publications co-authored with academic institutions. This variable had a positive but not significant effect when added to the model with the full set of variables.

5 Conclusion

This study examines the impact of scientific activities of firms on their stock market valuation using a dataset of firm-level information for US high technology companies, combined with scientific publication and patent data. Although scholars have recently started to conceptually and empirically address the determinants of scientific activities by firms, in particular disclosure, there is very little data for firm performance effects. Thanks to its design, our analysis provides an empirical contribution not only to the growing literature on boundary-spanning activities of firms, but also to more mainstream studies assessing the performance implications of R&D in firms.

Our analysis documents the positive impact of science-related indicators on a firm's market valuation beyond the effects of R&D and inventive outcome indicators. Our measures enabled us to differentiate between the profitability implications of performing research and a more active involvement in open science. While we cannot entirely rule out the possibility that we also capture research productivity effects, our findings suggest that active involvement in science, as reflected by disclosure in scientific journals, results in higher stock-market values. Furthermore, our results indicate in particular that the positive effects of such disclosure stem from scientific signaling to upstream stakeholders, which results in superior knowledge flows from these persons and institutions. In other words, publishing allows firms to become members of the scientific community and to establish formal and informal interactions, which in turn provide the former with access to state-of-the-art developments and research techniques. On the other hand, we found little evidence for a value-enhancing use of publications for appropriation purposes. Overall, our study provides support for the increasingly strong view among scholars that minimizing knowledge outflows may not be an optimal choice for firms and that they should instead strategically disclose knowledge.

Some heterogeneity concerning the impact of scientific activities was found. A negative but non-significant sign was found for ICT sectors. Instead, a strong positive relationship was observed in the instruments sectors. Therefore, appropriability conditions would seem to matter, which implies that R&D managers need to carefully consider the benefits and potential costs caused by related spillovers.

To obtain a better understanding of the mechanisms involved, future work could focus on contextual conditions that potentially moderate the relationship between scientific activities and the firm's market value. In conclusion, the strengths and weaknesses of the market value approach and corresponding empirical limitations should be kept in mind when interpreting results. Financial markets can be very volatile, and do not behave completely rationally. Thus, future research could complement our present study by using direct financial performance measures. This however, would entail dealing explicitly with the time lags between knowledge creation and commercial returns.

Acknowledgements

The authors thank Dirk Czarnitzki, Dominique Foray, Stuart Graham, Karin Hoisl, Stéphane Lhuillery, and Christian Sternitzke for valuable comments. The authors also thank audience members at conferences and university seminars in Torino, Mannheim, Barcelona, Bordeaux, Seville, Lisbon, Nice, and Madrid for helpful suggestions. The usual disclaimer applies. Markus Simeth acknowledges financial support from the Swiss National Science Foundation, grant number P2ELP1-151980.

References

- Aldieri L, Cincera M (2009) Geographic and technological spillovers within the Triad: Micro evidence from US patents. *Journal of Technology Transfer* 34(2): 196-211.
- Arrow, KJ (1962) *Economic welfare and the allocation of resources for innovation,*' in: HM Groves (ed.) *The rate and direction of inventive activity: Economic and social factors*. Princeton University Press, Princeton.
- Azoulay P (2002) Do pharmaceutical sales respond to scientific evidence? *Journal of Economics and Management Strategy* 11(4): 551-594.
- Belenzon S (2012) Cumulative innovation: Evidence from patent citations. *Economic Journal* 122(559): 265-285.
- Blundell R, Griffith R, Van Reenen J (1999) Market share, market value and innovation in a sample of British manufacturing firms. *Review of Economic Studies* 66(3): 529-554.
- Ceccagnoli M (2009) Appropriability, preemption, and firm performance. *Strategic Management Journal* 30(1): 81-98.
- Cockburn IM, Griliches Z (1988) Industry effects and appropriability measures in the stock market's valuation of R&D and patents. *American Economic Review* 78(2): 419-423.
- Cockburn IM, Henderson RM (1998) Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *The Journal of Industrial Economics* 46(2): 157-182.
- Cockburn IM, Henderson RM, Stern S (1999) Balancing Incentives: The tension between basic and applied research. NBER Working Paper 6882.
- Cockburn IM, Agrawal A, Bessen J, Graham JHS, Hall BH, MacGarvie M (2009). The NBER patent citation datafile update. <https://sites.google.com/site/patentdataprotect/Home>
- Cohen WM, Levinthal DA (1989) Innovation and learning: the two faces of R&D. *The Economic Journal* 99(397): 569-596.
- Cohen WM, Nelson RR, Walsh JP (2000). Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). NBER Working Paper No. 7552.
- Cohen WM, Nelson RR, Walsh JP (2002). Links and impacts: the influence of publish research on industrial R&D. *Management Science* 48 (1), 1-23.
- Czarnitzki, D, Hall BH, Oriani R (2006). Market Valuation of US and European Intellectual Property. In Bosworth, D. and E. Webster (eds), *The Management Of Intellectual Property*, Cheltenham, UK, Edward Elgar.
- De Fraja G (1993) Strategic spillovers in patent races. *International Journal of Industrial Organization* 11(1): 139-146.
- Della Malva A, Hussinger K (2012) Corporate science in the patent system: an analysis of the semiconductor technology. *Journal of Economic Behavior and Organization* 84(1): 118-135.
- Deng Z, Lev B, Narin F (1999) Science and Technology as predictors of stock performance. *Financial Analysts Journal* 55(3): 20-32.
- Fabrizio KR (2009) Absorptive capacity and the search for innovation. *Research Policy* 38(2): 255-267.
- Fleming L, Sorenson O (2004) Science as a map in technological search. *Strategic Management Journal* 25(8-9): 909-928.
- Gittelman M, Kogut B (2003) Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* 49(4): 366-382.
- Griliches Z (1981) Market value, R&D, and patents. *Economic Letters* 7(2): 183-187.

- Griliches, Z, Mairesse J (1984) Productivity and R&D at the Firm Level, NBER Chapters, in: R & D, Patents, and Productivity, pages 339-374, National Bureau of Economic Research.
- Hall BH (1993) The stock market's valuation of R&D investment during the 1980's. *American Economic Review* 83(2): 259-264.
- Hall BH (2000). Innovation and Market Value. In Barrell, Ray, Geoffrey Mason, and Mary O'Mahoney (eds.), *Productivity, Innovation and Economic Performance*, Cambridge: Cambridge University Press, pp. 175-198.
- Hall BH, Jaffe, AB, Trajtenberg, M (2001). The NBER Patent Citation data file: Lessons, Insights and Methodological Tools. NBER Working Paper 8498.
- Hall BH, Jaffe A, Trajtenberg M (2005) Market value and patent citations. *RAND Journal of Economics* 36(1): 16-38.
- Hall BH, Oriani R (2006) Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany, and Italy. *International Journal of Industrial Organization* 24(5): 971-993.
- Hall BH, Thoma G, Torrisi S (2007) The market value of patents and R&D: Evidence from European firms. NBER Working Paper 13426.
- Hall BH, Ziedonis RH (2001) The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979-1995. *RAND Journal of Economics* 32(1):101-128.
- Hicks D (1995) Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change* 4(2): 401-424.
- Jaffe AB (1986) Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review* 76(5): 984-1001.
- Kim J, Marschke G (2005) Labor mobility of scientists, technological diffusion, and the firm's patenting decision. *RAND Journal of Economics* 36(2): 298-317.
- Kinney AJ, Krebbers E, Vollmer SJ (2004) Publications from industry - Personal and Corporate Incentives. *Plant Physiology* 134(1): 11-15.
- Lanjouw JO, Schankerman M (2004) Patent quality and research productivity: Measuring innovation with multiple indicators. *Economic Journal* 114(495): 441-465.
- Levin RC, Klevorick AK, Nelson RR, Winter SG (1987). Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity* 1987(3), 783-831.
- Stuart TE, Liu C (2014) Positions and rewards: The allocation of resources within a science-based entrepreneurial firm. *Research Policy* 43(7), 1134-1143.
- Marco AC (2007) The dynamics of patent citations. *Economic Letters* 94(2): 290-296.
- McGahan AM, Silverman BS (2006) Profiting from technological innovation by others: The effect of competitor patenting on firm value. *Research Policy* 35(8): 1222-1242.
- Narin F, Hamilton KS, Olivestro D (1997). The increasing linkage between US technology and public science. *Research Policy* 26, 317-330.
- Nelson RR (1959) The simple economics of basic scientific research. *The Journal of Political Economy* 67(3): 297-306.
- OECD (2011) ISIC Rev. 3 Technology intensity definition: Classification of manufacturing industries into categories based on R&D intensities. *OECD Directorate for Science, Technology & Industry*: [<http://www.oecd.org/sti/industryandglobalisation/48350231.pdf>]; Last access: 17th October 2012.
- Parchomovsky G (2000) Publish or perish. *Michigan Law Review* 98(4): 926-952.
- Penin J (2007) Open knowledge disclosure: an overview of the evidence and economic motivations. *Journal of Economic Surveys* 21(2): 326-347.

- Pisano GP, 2006. *Science Business. The promise, the reality, and the future of biotech.* Harward Business Review Press, Boston.
- Polidoro F, Theeke M (2012) Getting competition down to a science: The effects of technological competition on firms' scientific publications. *Organization Science* 23(4): 1135-1153.
- Raffo J, Lhuillery, S (2009) How to play the "Names Game": Patent retrieval comparing different heuristics. *Research Policy* 38(10): 1617-1627.
- Rosenberg, N (1990) Why firms do basic research (with their own money). *Research Policy* 19(2): 165-174.
- Sauermann H, Roach M (2013) Not all scientists pay to be scientists: PhD's preferences for publishing in industrial employment. *Research Policy* 43(1), 32-47.
- Simeth M, Raffo JD (2013) What makes companies pursue an Open Science strategy? *Research Policy* 42(9): 1531–1543.
- Stern S (2004) Do scientists pay to be scientists? *Management Science* 50(6): 835-853.
- Trajtenberg M, Henderson R, Jaffe A (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology* 5 (1): 19-50.

TABLES AND FIGURES

Table 1: Observations according to included sectors

Meta-Sector	SIC included	Firm-year observations
Biotechnology & Pharmaceuticals	2834, 2835, 2836	1 602
Chemicals	2800, 2810, 2820, 2821, 2833	389
Information and Communication Technologies (ICT)	3570, 3571, 3572, 3575, 3576, 3577, 3578, 3579, 3661, 3663, 3669, 3670, 3672, 3674, 3677, 3678, 3679, 4812, 4813, 4822	4 373
Aircraft & Aerospace	3721, 3724, 3728	166
Navigation, Scientific, Medical, and Optical instruments	3812, 3822, 3823, 3824, 3825, 3826, 3827, 3829, 3841, 3842, 3843, 3844, 3845, 3851, 3861	3 390
Sum		9 920

Table 2: Summary statistics

Variable	N	Mean	Std. Dev.	Median	Min	Max
TOBIN'S Q	9920	2.65	2.39	1.89	0.25	23.02
A (BOOK VALUE)	9920	1548.56	7172.64	94.33	0.55	273007.30
R&D (FLOW)	9920	104.29	468.25	8.36	0.00	12942.19
R&D (STOCK)	9920	491.47	2082.16	39.96	0.00	41408.40
R&D / A	9920	0.61	0.71	0.39	0.00	7.42
PAT (STOCK)	9920	114.33	559.85	5.57	0.00	9969.30
PAT / R&D	9920	0.31	0.46	0.15	0.00	3.68
SCIPAT / PAT	9920	0.27	0.31	0.17	0.00	1.00
FWDCIT / PAT	9920	5.59	7.50	4.38	0.00	198.65
PUB (STOCK)	9920	69.54	347.81	1.77	0.00	6440.20
PUB / R&D	9920	0.15	0.29	0.04	0.00	2.85
TOPPUB / PUB	9920	0.20	0.29	0.00	0.00	1.00
SALES	9920	1229.41	5248.79	79.41	0.50	88784.06
EMPLOYEES	9738	4076.93	14324.04	340.00	10.00	238000.00

Monetary amounts in Mio USD (2005, GDP deflated)

Table 3: Regression outputs

LOG TOBIN'S Q	(1)	(2)		(3)	(4)	(5)	(6)		(7)
	OLS	NLLS		NLLS	NLLS	NLLS	NLLS		OLS
	Coeff (SE)	Coeff (SE)	Semi-elast.	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Semi-elast.	Coeff (SE)
R&D/A	0.048*** (0.011)	0.055*** (0.014)	0.051	0.055*** (0.015)	0.052*** (0.014)	0.051*** (0.014)	0.050*** (0.015)	0.045	0.042*** (0.011)
PAT/R&D	0.083*** (0.021)	0.097*** (0.025)	0.090	0.071*** (0.026)	0.083*** (0.025)	0.097*** (0.025)	0.072*** (0.026)	0.064	0.066*** (0.021)
PUB/R&D	0.095*** (0.032)	0.110*** (0.039)	0.102	0.108*** (0.039)	0.091** (0.039)	0.066* (0.038)	0.065* (0.039)	0.057	0.056* (0.032)
FWDCIT/PAT				0.007*** (0.002)			0.006*** (0.002)	0.005	0.004** (0.002)
SCIPAT/PAT					0.124*** (0.035)		0.046 (0.037)	0.041	0.052 (0.032)
TOPPUB/PUB						0.168*** (0.040)	0.142*** (0.041)	0.126	0.125*** (0.034)
Firm-Year observations	9,920	9,920		9,920	9,920	9,920	9,920		9,920
Firm-IDs (cluster)	1.739	1.739		1.739	1.739	1.739	1.739		1.739
R ²	0.129	0.130		0.136	0.134	0.136	0.141		0.141

Standard errors (clustered by firm) reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

All regression models contain year and SIC-4 digit industry dummies. The semi-elasticities reported in the table are computed at the mean values of the variables.

The semi-elasticities of the added variables in columns (3)-(5) are as follows: *FWDCIT/PAT* 0.007, *SCIPAT/PAT* 0.113, and *TOPPUB/PUB* 0.153.

Table 4: Heterogeneity by sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	BIO-PHARM		ICT		INST		CHEM	
LOG TOBIN'S Q	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS
	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)
R&D/A	0.126*** (0.040)	0.108*** (0.041)	0.022 (0.017)	0.022 (0.018)	0.086*** (0.029)	0.075** (0.030)	0.395* (0.208)	0.389* (0.206)
PAT/R&D	-0.040 (0.067)	-0.028 (0.075)	0.142*** (0.049)	0.110** (0.048)	0.103*** (0.036)	0.079** (0.038)	0.067 (0.115)	-0.010 (0.094)
PUB/R&D	0.133 (0.088)	0.045 (0.086)	-0.009 (0.049)	-0.014 (0.050)	0.218*** (0.075)	0.149** (0.075)	0.115 (0.169)	0.058 (0.140)
FWDCIT/PAT		0.006 (0.007)		0.003 (0.002)		0.009* (0.005)		0.022 (0.020)
SCIPAT/PAT		-0.037 (0.074)		0.129** (0.058)		0.021 (0.070)		-0.098 (0.141)
TOPPUB/PUB		0.225** (0.090)		-0.015 (0.054)		0.243*** (0.077)		0.026 (0.137)
Firm-Year observations	1,602	1,602	4,373	4,373	3,390	3,390	389	389
Firm-IDs (cluster)	331	331	762	762	565	565	57	57
R ²	0.071	0.088	0.120	0.127	0.129	0.156	0.136	0.150

Standard errors (clustered by firm) reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

All regression models contain year and SIC-4 digit industry dummies.

Regression models represent baseline and full model specification for meta-sectors biotechnology and pharmaceuticals (BIO-PHARM) in columns (1) and (2), for information and communication technologies (ICT) in (3) and (4), for instruments (INST) in (5) and (6), and for Chemicals (CHEM) in (7) and (8). See Table 1 for information on SIC codes included in the meta-sectors. The regression results for the Aircraft sector are omitted in this table due to the very low number of observations. The semi-elasticities of the publication-based measures are as follows: *PUB/R&D*: 0.039 in column (2), -0.014 in (4), 0.125 in (6), 0.058 in (8). *TOPPUB/PUB*: 0.195 in column (2), -0.014 in (4), 0.204 in (6), and 0.022 in (8).

Figure 1a: Average publication & patent outputs

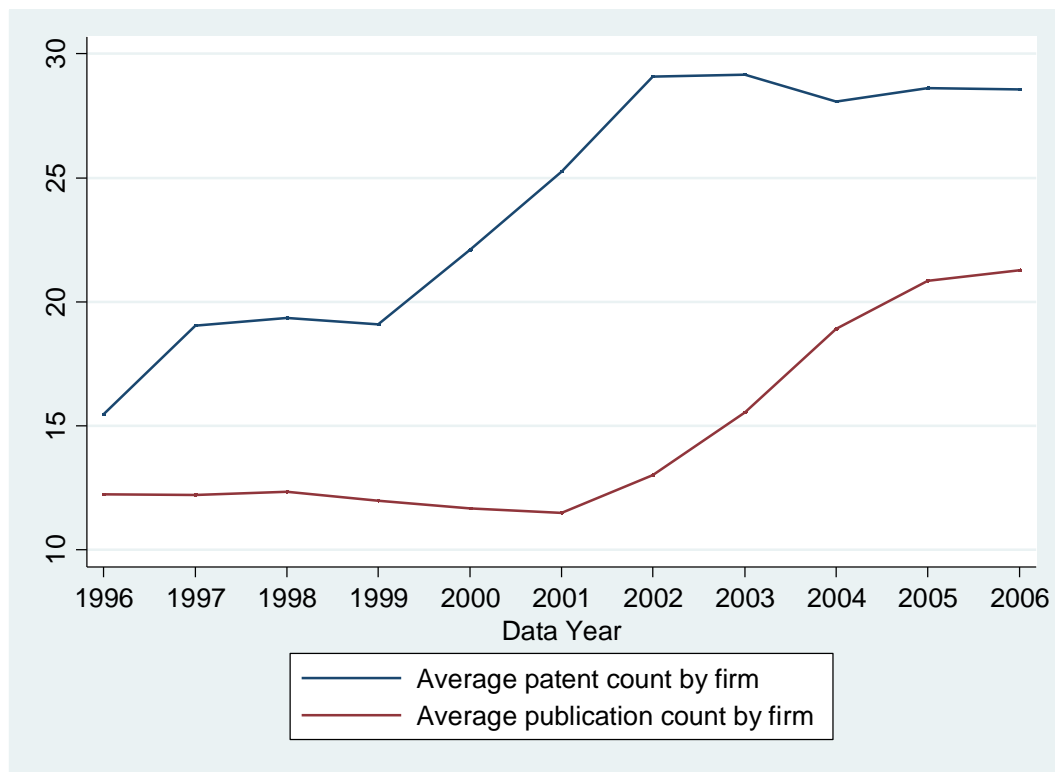
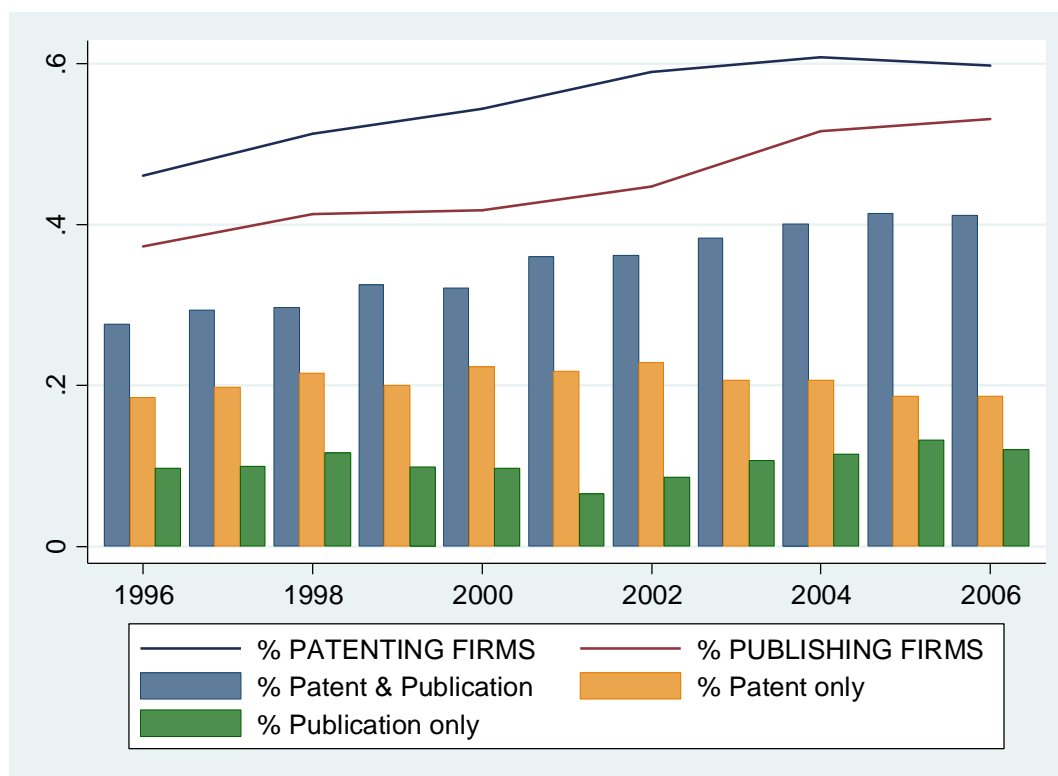


Figure 1b: Share publishing and patenting firms



APPENDIX

Appendix A1: Matching firm-level data with publication data

In the absence of any direct linking identifiers, the firm-level data from COMPUSTAT and data from SCOPUS have to be combined manually using the firm name (and potential abbreviations) and the affiliation field in the publication database. The firm names – as listed in the *COMPUSTAT* database – contain legal identifiers and other elements which may lead to too few matches (“false negatives”). Therefore, as a general rule, we reduced the firm names to the most distinctive name components, removing frequent name elements and legal identifiers in order to achieve high recall rates. Frequent examples included the words “Pharmaceuticals”, “Technologies”, or “International” which typically have little identifying power and lead to false negatives when the publication authors do not mention the full official firm name. Moreover, although the Scopus engine is not sensitive to name capitalisation and language-specific characters, we harmonized firm names correspondingly since this is a useful step for post-download processing. It has to be considered that reducing the firm names to a minimum may have increased the number of false positive search results. A common source of false positives are, for example, university units or independent firms with a common name component. Thanks to the manual downloading process, all search results were inspected directly and in the case of ambiguity, a complementary web search was used to select only eligible publications. Finally, we performed post-download verification checks by transforming the affiliation information of the different authors of a particular publication into a “long” format, harmonizing the affiliation field with regard to name capitalizations and language-specific characters, and applying queries to detect whether the exact field that created the link between the firm and publication record simultaneously contained an element pointing to an academic institution. In these rare cases, the publication records were manually checked.

Example: ADHEREX TECHNOLOGIES INC

This name contains with “Technologies” and “INC” two name components that potentially lead to too few search results. The identifier INC was therefore directly removed. Whereas the search using “ADHEREX TECHNOLOGIES” led to 21 publications in the period between 1996 and 2010, using only “ADHEREX” returned one additional publication. After manual inspection of the search results, we concluded that all 22 publications were eligible and we subsequently downloaded the corresponding information. Query: *AFFILORG (“ADHEREX”) AND PUBYEAR > 1995 AND PUBYEAR < 2007*

Affiliation		
<input type="checkbox"/> Adherex Technologies Inc. (15)	<input type="checkbox"/> Cancer and Leukemia Group B Statistical Center (1)	<input type="checkbox"/> The Johns Hopkins School of Medicine (1)
<input type="checkbox"/> Adherex Technologies Inc. (6)	<input type="checkbox"/> Piedmont Research Center (1)	<input type="checkbox"/> Scripps Florida (1)
<input type="checkbox"/> McGill University (3)	<input type="checkbox"/> Southern Europe New Drug Organization Foundation (1)	<input type="checkbox"/> The Research Triangle Park (1)
<input type="checkbox"/> Duke University School of Medicine (3)	<input type="checkbox"/> Memorial Regional Trauma Center (1)	<input type="checkbox"/> LSU Health Sciences Center - Shreveport (1)
<input type="checkbox"/> GlaxoSmithKline (3)	<input type="checkbox"/> Spector Consulting Services (1)	<input type="checkbox"/> Fondazione IRCCS Istituto Nazionale dei Tumori, Milan (1)
<input type="checkbox"/> Case Western Reserve University (2)	<input type="checkbox"/> Kansai Laboratory (1)	<input type="checkbox"/> Memorial Sloan-Kettering Cancer Center (1)
<input type="checkbox"/> Duke University (2)	<input type="checkbox"/> Critical Thinking Co. (1)	<input type="checkbox"/> University of Hawaii at Manoa (1)
<input type="checkbox"/> King's College London (2)	<input type="checkbox"/> Tuskuba Research Laboratories (1)	<input type="checkbox"/> Western General Hospital (1)
<input type="checkbox"/> University of Texas M. D. Anderson Cancer Center (2)	<input type="checkbox"/> Instituto Oncologico della Svizzera Italiana (1)	<input type="checkbox"/> VA Medical Center (1)
<input type="checkbox"/> Royal Victoria Hospital Barrie (2)	<input type="checkbox"/> Adherex, Inc. (1)	<input type="checkbox"/> Istituto Nazionale Tumori Regina Elena (1)

Limit to Exclude

Appendix A2: Subsidiary publications

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS
LOG TOBIN'S Q	SUB	NO SUB	SUB	NO SUB	SUB	NO SUB	SUB	NO SUB
	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)
R&D/A	0.065* (0.033)	0.065* (0.033)	0.080* (0.044)	0.082* (0.043)	0.342*** (0.088)	0.340*** (0.087)	0.117*** (0.029)	0.116*** (0.029)
PAT/R&D	0.057 (0.071)	0.056 (0.071)	-0.049 (0.093)	-0.051 (0.094)	0.140** (0.061)	0.139** (0.061)	0.052 (0.041)	0.051 (0.041)
PUB/R&D	0.128 (0.092)	0.136 (0.097)	-0.025 (0.076)	-0.002 (0.104)	0.064 (0.132)	-0.020 (0.083)	0.104* (0.062)	0.106 (0.066)
Firm-Year observations	1,126	1,126	706	706	798	798	2,630	2,630
Firm-IDs (cluster)	186	186	160	160	145	145	491	491
R ²	0.123	0.124	0.056	0.055	0.144	0.143	0.069	0.069

Standard errors (clustered by firm) reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

All models contain year dummies.

Column (1a) and (1b): SIC 3674 (Semiconductors); (2a) and (2b): SIC 2835 & 2836 (In vitro & in vivo diagnostic substance; Biological products); (3a) and (3b): SIC 3841 & 3842 (Surgical & Medical Instruments, Orthopedic and surgical appliances); (4a) and (4b): All sectors combined from columns (1)-(3).

Appendix A3: Definition of top-impact journals

General information

The distinction as to whether a firm publication was of a high quality was made using journal-impact factor information. The data source for journal impact factors was the SCIMAGO classification, which contains the *SCImago Journal Rank (SJR)* indicator based on the records in the Elsevier's SCOPUS database. The SCIMAGO ranking is propagated by Elsevier as an official indicator for SCOPUS. Since the journal rankings are dynamic but change in reality slowly over time, we chose the reference year 2000 which is the centre year in our sample period.

URL: <http://www.scimagojr.com> and <http://www.journalmetrics.com>

Construction of indicator

In order to normalize the indicator for field effects potentially not entirely considered in the SJR indicator, we created Journal Impact factor rankings specifically for the meta-fields Life Sciences, Physical Sciences, Health Sciences, General journals, Social Sciences. We created our top-impact journal indicator based on the distribution of the SJR indicator within the five meta-disciplines, and classified the top 10% of all journal sources as *top-impact journals*. The meta-fields contain the following number of journals:

- * LIFE SCIENCES: 4,646 sources
- * PHYSICAL SCIENCES: 8,916 sources
- * HEALTH SCIENCES: 8,483 sources
- * GENERAL JOURNALS: 95 sources
- * SOCIAL SCIENCES: 6,687 sources

Examples of journal classification in the Innovation Research field

To illustrate the outcome of the classification for journals known in the Management field, some examples are presented below:

Classified as Top Impact Journals: Management Science, Journal of Industrial Economics, Journal of Economics and Management Strategy, RAND Journal of Economics; Journal of Economic Behaviour and Organization, Research Policy

Not classified as Top Impact Journals: Industrial and Corporate Change, Journal of Technology Transfer, Technovation, Industry and Innovation, Small Business Economics

Appendix A4: Robustness tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	FE	R&D Filter	LOC D	ORIGPAT	ORIGPAT	Controls	Controls	Controls	NBER	PUB/PAT	PUB/PAT	ACADPUB
LOG TOBIN'S Q	WITHIN	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS	NLLS
	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)
R&D/A	0.026*** (0.009)	0.078*** (0.014)	0.046*** (0.014)	0.056*** (0.015)	0.051*** (0.015)	0.051*** (0.015)	0.058*** (0.015)	0.062*** (0.015)	0.059*** (0.017)	0.062*** (0.017)	0.056*** (0.018)	0.050*** (0.015)
PAT/R&D	0.073*** (0.019)	0.074*** (0.026)	0.076*** (0.025)	0.084*** (0.025)	0.069*** (0.026)	0.074*** (0.028)	0.083*** (0.026)	0.064*** (0.025)	0.075** (0.030)	0.057** (0.028)	0.075** (0.030)	0.071*** (0.026)
PUB/R&D	0.058** (0.027)	0.044 (0.034)	0.077* (0.040)	0.102*** (0.039)	0.064 (0.039)	0.072* (0.040)	0.036 (0.039)	0.042 (0.039)	0.113*** (0.043)			0.060 (0.039)
PUB/PAT										0.269** (0.111)	0.176 (0.120)	
FWDCIT/PAT	0.001 (0.001)	0.005*** (0.002)	0.005** (0.002)		0.005*** (0.002)	0.006*** (0.002)	0.007*** (0.002)	0.006*** (0.002)			0.006** (0.003)	0.005*** (0.002)
SCIPAT/PAT	-0.039 (0.027)	0.016 (0.034)	0.037 (0.037)		0.036 (0.039)	0.041 (0.040)	0.054 (0.040)	0.050 (0.039)			0.055 (0.042)	0.043 (0.038)
TOPPUB/PUB	0.033 (0.027)	0.116*** (0.037)	0.135*** (0.041)		0.139*** (0.041)	0.155*** (0.044)	0.173*** (0.043)	0.155*** (0.042)			0.215*** (0.049)	0.111** (0.046)
ORIGPAT/PAT				0.097*** (0.032)	0.038 (0.034)							
ACADPUB/PUB												0.052 (0.039)
Firm-Year observations	9,920	11,420	9,920	9,920	9,920	9,154	9,920	9,920	7,392	7,407	7,407	9,920
Firm-IDs (cluster)	1,739	1998	1739	1,739	1,739	1593	1,739	1,739	1,575	1,312	1,312	1,739
R ²	0.089	0.158	0.154	0.132	0.141	0.152	0.112	0.120	0.140	0.140	0.156	0.142

Standard errors (clustered by firm) reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Models contain year and SIC-4 digit industry dummies.

Column (1): FE-"Within" estimator. (2): Filter that excludes R&D/Sales > 1 is deleted. (3): Location dummies (US State level) are added. (4) and (5): Patent originality measure (ORIGPAT/PAT) is introduced. The threshold for considering a patent as "original" has been set at the mean originality value of all patents filed by the sample firms. (6): Sales and Sales growth are included as control variables (see Belenzon 2011). The sample size is reduced since Sales growth requires observation of t-1. (7) contains sector-level patent and publication intensities (computed at NACE-3 level) as additional control variables and "metasector" industry controls. (8) contains a control variable that captures the share of firm's profits on industry profits (at SIC-3 digit level). (9): Patentcounts from the combined Compustat-NBER dataset is used instead of the variable obtained by our own matching. Sample limited to the period 1996-2002 since the NBER patent counts are increasingly incomplete for more recent years. (10) and (11): Alternative consideration of publications as ratio of publication stocks to patent stocks, for subsample of patenting firms. (12): Share co-authored publication stock with academic institutions is introduced.