Citation for published version:
Zahringer, K, Kolympiris, C & Kalaitzandonakes, N 2017, 'Academic knowledge quality differentials and the quality of firm innovation', *Industrial and Corporate Change*, vol. 26, no. 5, pp. 821-844.
https://doi.org/10.1093/icc/dtw050

DOI:
10.1093/icc/dtw050

Publication date:
2017

Document Version
Peer reviewed version

Academic knowledge quality differentials and the quality of firm innovation

Kenneth Zahringer
Research Fellow
Economics and Management of Agrobiotechnology Center
University of Missouri
kaz998@mail.missouri.edu
Phone: (573) 884 1602

Christos Kolympiris*
Associate Professor
School of Management, University of Bath
c.kolympiris@bath.ac.uk
Phone: +44 01225384972

Nicholas Kalaitzandonakes
MSMC Endowed Professor of Agribusiness Strategy
Director, Economics and Management of Agrobiotechnology Center
University of Missouri
kalaitzandonakesn@missouri.edu
Phone: (573) 882 0143

This paper examines whether knowledge from academic research of varying quality can lead to differential quality innovation at the firm level via a knowledge spillover process – without the firm conducting the research and producing the knowledge but by using it as an external input to its innovation process. Based on econometric analysis of data reflecting patent activity of emerging entrepreneurial life sciences firms and thousands of academic publications we find that higher quality academic science associates with higher quality industrial innovation. Importantly, we document that the degree that a given firm conducts basic science moderates this relationship. Our work develops a novel theoretical and empirical link between academic knowledge quality and industrial innovation and identifies the contingencies that allow this link to materialize.

* Corresponding author

Disclaimer: This material is based upon work supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under Agreement No. 2008-38420-18747. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Agriculture.

1. Introduction

Recognizing that the locus of innovation often resides outside an organization's boundaries (Cohen and Levinthal, 1990), firms in knowledge intensive industries regularly use academic science as an external input to their innovation process (Cassiman et al., 2010; Cassiman et al., 2008). Importantly, academic science is heterogeneous in quality; some works provide more insights than others, present more novel techniques, are better executed and generally present a richer knowledge pool that firms can draw upon (Gittelman, 2007; Jong and Slavova, 2014; Nonaka, 1994). Putting those two phenomena together begs the question: does higher quality academic science as an external input to the innovation process translate to higher quality firm innovation?

The extant literature has not analyzed this question in any significant depth as the focus has been on examining the linkages between academic science and the quality of industrial innovation. Subsequently, that focus has shifted attention away from the heterogeneity of academic science quality. Cassiman et al. (2008), Fleming and Sorenson (2004), Nagaoka (2007) and Sorenson and Fleming (2004) all conclude that knowledge-intensive firms which use academic research as an external input to the innovation process improve the quality of their innovations. They do not, however, take into account the fact that academic research is heterogeneous in quality. That is, they reach their important insights by linking firm patent quality measures with the number of academic publications that are cited in the firm patents but are not authored by the focal firm. However, the only quality indicator these studies use is a separation of journal articles with an impact factor from the full list of prior art references[1]. As such, they treat all articles (as long as they are published in a journal which has an impact factor)

[1] Nagaoka (2007) does not make this separation relying instead on the count of citations to academic papers.

as being of similar quality. Accordingly, they mask potentially substantial differences in academic science quality. These differences may manifest both at an aggregate level (i.e. overall journal quality, as reflected in impact factors) and at a micro level (i.e. article quality, as reflected in number of citations garnered by a given article irrespective of the journal in which it is published). These are important considerations given the wide variation in quality not only among journals but also within journals, with certain contributions being of higher quality than others.

Against this background and motivated by studies linking knowledge quality measures (e.g. citation weighted publications) of *in-house* basic research with innovation outcomes (Almeida et al., 2011; Belenzon and Patacconi, 2014; Gittelman and Kogut, 2003; Jong and Slavova, 2014; Subramanian et al., 2013), in this paper we bring academic knowledge quality, measured with impact factors and article citation counts, as an external input to industrial innovation to the forefront. We do so both conceptually and empirically. After we review what is known about the nature of academic science, the derived knowledge, and the process that firms use to innovate we discuss the mechanisms that can translate higher science quality to higher quality of industrial innovation. Importantly, given the cognitive difficulties that some firms may have in assimilating such knowledge (Kealey and Ricketts, 2014) we then uncover the boundary conditions under which such relationship holds. In line with the literature on the benefits afforded to firms conducting basic science (Gambardella, 1992; Rosenberg, 1990) and with the literature highlighting the positive effects of developing higher absorptive capacity (Cohen and Levinthal, 1990; Fabrizio, 2009; Lane and Lubatkin, 1998) we find that only firms conducting and publishing in-house basic research are able to produce higher quality innovations by using higher quality academic science as an external input to the innovation process.

In the empirical part of the paper we measure knowledge from academic research with articles published in academic journals. We do so based on robust findings which document that journal articles are a common vehicle by which industrial actors tap into academic knowledge (Cohen et al., 2002) and represent a reliable measure of knowledge flows from the public sector (Roach and Cohen, 2013). To measure the quality of academic research we use two separate measures. The first reflects the number of citations received by each journal article referenced in each patent. The second measures the impact factor of the journals in which these articles are published. To explicitly test whether higher quality knowledge from academic research leads to higher quality of innovations at the firm level we build a number of econometric specifications that link patent quality with the number of published articles that are cited in the focal patents and the knowledge quality variables we described above.

We conduct the analysis within a sample of emerging entrepreneurial life sciences firms. We focus our attention in the area of the life sciences as academic research in this area tends to be both a starting point of industrial development (Balconi et al., 2010) and to have immediate industrial applications (Argyres and Liebeskind, 1998; Narin et al., 1997). In line with a long stream of research, we measure the quality of innovation through patent quality (Hagedoorn and Cloodt, 2003; Lanjouw and Schankerman, 1999) which we approximate with the number of times a patent has been cited by later patents (i.e. forward citations) (Fischer and Leidinger, 2014; Harhoff et al., 1999; Lerner, 1994; Sneed and Johnson, 2009).

Our choice to concentrate on innovations produced by emerging, mostly young entrepreneurial firms in the life sciences is aligned with the use of academic science as an external input.[2] Typically, such firms have limited financial resources and often rely on

[2] We opt to measure technological innovation with patents based on the evidence that in high technology industries patents are a reliable innovation proxy despite their shortcomings (Acs et al., 2002; Igami, 2013). These

knowledge spillovers (Khilji et al., 2006; Kolympiris and Kalaitzandonakes, 2013) and knowledge from academic research (Cohen et al., 2002; Laursen and Salter, 2004) to fuel their innovation. As such, if the differential quality of knowledge from academic research has a differential impact on industrial innovation, we expect to observe such relationship in the case of emerging entrepreneurial life sciences firms. Note that these types of firms differ in the degree they conduct and publish basic research which then allows us to uncover the boundary conditions discussed above.

For the analysis we construct and merge 3 different datasets. The first identifies and describes all life-sciences grant recipients of the Small Business Innovation Research (SBIR) program and it is provided by commercial vendor InKnowVation. As we explain below in detail, the SBIR program awards grants to the types of firms we are interested in - emerging small innovative firms. The second dataset measures the patent activity of those firms across time; it was assembled by sourcing proprietary data from Thomson Innovation. The third dataset contains detailed information on all the publications cited in the focal patents as well as on the publication activity of the sample firms. It is constructed via intensive programming and manual cleanups utilizing data from the bibliographic database SCOPUS.

Understanding the relationship between quality of knowledge from academic research and industrial innovation in small entrepreneurial firms is important in its own right. Small and young firms, especially in knowledge intensive industries such as the life sciences, contribute significantly to innovation, job creation, wealth generation and the like (Cumming et al., 2014; Haltiwanger et al., 2013; Henrekson and Johansson, 2010; Lawless, 2014; Neumark et al.,

shortcoming may arise because innovative firms may not patent for strategic reasons (Teece, 1986) or patents may relate more to invention rather than to innovation (Moser, 2005). Specifically for the life sciences the use of patents as an innovation proxy is particularly suitable because of the strong links between innovation and patents in that industry (Arundel and Kabla, 1998).

2011).[3] Thus, understanding the factors that drive improvements in the quality of the innovation process in those firms should be important for social welfare.

The rest of the paper is organized as follows: In section 2 we discuss the notion of academic knowledge quality and we summarize the mechanisms that can connect higher academic quality to higher innovation quality as well as the contingencies that allow such links to materialize. In section 3 we describe our empirical models. In section 4 we discuss our data source, sample development and variable construction. In section 5 we present our results and we conclude in section 6.

2. Published academic research, its quality and industrial innovation

Academic research quality is an encompassing term that incorporates notions relating, among others, to academic work that has more societal impact, affords higher recognition among peers, exerts a stronger influence on the advance of knowledge, it is more important, makes stronger contribution to the body of knowledge and inspires further research (e.g. Aksnes and Rip, 2009; Jensen et al., 2003; Owen-Smith and Powell, 2003; Stremersch et al., 2007). Largely because of its encompassing nature, defining research quality with precision is an inherently difficult task as different actors may assign different weights in various components of quality. For instance, some may place more emphasis on societal impact, others may look for novelty, while others may evaluate more favorably work that forms the foundation for follow up studies. While such differences, when they exist, obstruct a comprehensive and commonly agreed upon definition, they do not pose a prohibitive obstacle on approximating the quality of a given piece of academic work. That holds largely because secondary data can represent most elements of research quality simultaneously. That is, journal rankings and impact factors, the number of times an article has

[3] Nightingale and Coad (2014) present a thorough account of the contribution of emerging firms to economic growth as well as on the conditions under which this relationship does not hold.

been cited by follow up work (forward citations) and other widely used indicators, despite their shortcomings, generally approximate research quality and offer two main advantages (e.g. Azoulay et al., 2014; Thursby and Kemp, 2002). First, they correspond well intuitively with the notions of recognition, impact and the like; to give an example, one can easily conceptualize forward citations as a measure of the degree that a given piece of academic research acts as foundation for follow up work. Second, and perhaps more fundamentally, they also tend to correlate with the assessment of research quality as revealed by judgements made by scientists (Aksnes, 2003, 2006; Aksnes and Rip, 2009). This latter observation complements existing findings suggesting that *ex ante* scores provided by scientists correlate with realized research quality (Li and Agha, 2015). This is particularly important because not all scientists adopt the same definition of quality, yet they generally find works that score high in the secondary measures in question to be of superior quality.

The next inquiry is if/how academic knowledge of higher quality can lead, via a knowledge spillover process, to higher quality industrial innovation. As already introduced, existing works pick up the issue only indirectly since, when explaining the quality of industrial innovation, they treat all articles (as long as they are published in a journal having an impact factor) as being of similar quality (Cassiman et al., 2008; Fleming and Sorenson, 2004).[4] Such empirical choice is in line with the theoretical framework of those contributions in which knowledge quality is not central as an input to firm innovation.[5] Alternatively, we propose that knowledge quality merits separate attention.

[4] Some works, in the context of in-house (academic) science production, acknowledge that knowledge can be heterogeneous in terms of its quality and therefore also in terms of potential relevance for industrial innovation (Gittelman, 2007; Jong and Slavova, 2014; Nonaka, 1994).

[5] Sorenson and Fleming (2004), without adjusting journal articles for underlying knowledge quality and without focusing on firm patents, attribute the positive association they find between forward citations and science intensity of a given patent to the rapid diffusion of knowledge that publications (as part of the prior art) bring about. Such explanation is indeed possible but it applies only when the dependent variable is the number of forward citations. In

Accordingly, in what follows we first discuss how academic science can contribute to industrial innovation at the firm level and then approach the question of how/if academic knowledge of higher quality can lead, via a knowledge spillover process, to higher quality of industrial innovation.

Following the dynamic theory of organizational knowledge creation (Nonaka, 1991, 1994) the interplay between tacit and explicit knowledge is what drives innovation within organizations. When tacit knowledge, defined as knowledge that is difficult to be fully articulated by its possessor and it is held in “people’s heads and bodies” (Leonard and Sensiper, 1998; Polanyi, 1975) is transformed to explicit knowledge, and vice versa, organizations innovate (Nonaka, 1991, 1994). This happens via articulation in which tacit knowledge is converted to codified knowledge and via internalization in which codified knowledge is used as an input to strengthen one’s knowledge base (Nonaka, 1991). When such processes occur, a spiral of knowledge is eventually created and organizations innovate. Leonard and Sensiper (1998) discuss how the exchange between tacit and explicit knowledge is crucial via three main routes. First, it helps in problem solving as it allows its possessor to identify patterns and reach solutions more effectively. Second, it allows for a better framing of problems which does not rely on established routines but rather on intuition and creative thinking. Third, it promotes an enhanced understanding of how processes work, and as such it allows for more accurate predictions.

---

unreported results where we employed patent family size as an alternative measure of patent quality we reach conclusions similar to those derived from the baseline estimates we present in Table 3. As such, we expect to indeed be measuring patent quality. Note that in analyzing forward citations of European patents Cassiman et al. (2008) had an interpretation of their dependent variable similar to ours. We acknowledge though, that the conclusions of Cassiman et al. (2008) are not directly comparable to our study given a) the significant differences in the patent offices between EU and the US, hence the differences in the content of patents (de la Potterie, 2011) and b) the fact that Cassiman et al. (2008) used only European patents of 79 Flemish firms.

When academic science is not produced in-house but it is used as an external input to the innovation process we expect it to contribute to firm innovation via the mechanism of internalization --- when codified knowledge enhances the knowledge base of a given firm (Nonaka, 1991). When it is not conducted internally, published academic science represents a source of codified knowledge for the focal firm, a source that can serve as an input to enrich a firm's knowledge base.[6] This may happen if external codified knowledge becomes an example to follow or a knowledge input when firms codify their internal tacit knowledge and, hence, increase the clarity of their ideas (Zollo and Winter, 2002). Similarly, external codified knowledge can boost a firm's knowledge base by facilitating the exchange of tacit and codified knowledge. For instance, in reverse engineering firm employees may use published academic (codified) work and build their own tacit knowledge based on it; the end result is the codification of tacit firm knowledge in the form of innovation. Such a process highlights the exchange between tacit and codified knowledge: codified external academic knowledge becomes an input in the generation of firm tacit knowledge and then is transformed into codified firm knowledge.

A second mechanism through which academic knowledge as an external input can boost firm innovation is when it minimizes the search costs involved in most innovation processes. When firms conduct (academic) research in-house they generate tacit knowledge in the process but they also incur the costs associated with the discovery process. This mechanism has been described in previous works (e.g. Hicks, 1995) but, by definition, does not apply when external academic

[6] Indeed, in the life sciences academic knowledge consistently fuels industrial progress (Argyres and Liebeskind, 1998; Cockburn and Henderson, 1998; Liebeskind et al., 1996) and academic life science research is often the cornerstone for subsequent commercial developments (Balconi et al., 2010). As the lines between basic and applied research are often blurry, academic research can have nearly immediate commercial applications (Argyres and Liebeskind, 1998; Narin et al., 1997). Accordingly, industrial players in the life sciences can draw upon knowledge emanating from academic research to advance their immediate innovative efforts (McMillan et al., 2000) in order to generate new projects and complete existing projects (Cohen et al., 2002). To illustrate, even before the boom in the life sciences industry, one out of three medical products and drugs produced from 1986 to 1994 in the US would not have been possible had it not been for academic research (Mansfield, 1998).

science is used as an external input to firm innovation. Instead, academic science, conceptualized as a map that describes the technological landscape (Fleming and Sorenson, 2004), can provide, as an external input, a better understanding of the technological landscape at a smaller cost. This holds insofar as it allows actors to in essence leverage (part of) the advantages of possessing tacit knowledge without having to incur the search costs and experience that lead to it. As Fleming and Sorenson (2004) put it, "science can tell inventors how to avoid wasted effort", which is a major advantage of possessing tacit knowledge. To be clear, using academic science as an external input does not eliminate all search costs because the process leading to the published outcome and the associated tacit knowledge are not revealed.

When, though, academic knowledge takes the form of published articles such search costs could indeed lessen as the publication process weeds out less valuable knowledge sources without the firm having to perform the task and incur the costs. That is, because what is typically reported in published articles are the most noteworthy new findings and scholarly contributions journal articles convey the latest academic knowledge that most merits attention and hence better describes the technological landscape without the need of personal contacts (Sorenson and Fleming, 2004). As a result, the process of publication serves as a filter of works that could be relevant as an external input to the innovation process. Additionally, published articles are also more relevant as a form of external academic knowledge when considering the mechanism of internalization discussed above. For instance, published articles allow the receiver of knowledge to learn from many sources (what Gray and Meister (2004) term published knowledge sourcing). It follows that firms may combine insights and may also have more examples to work with when they codify their internal tacit knowledge.

The key factor underpinning the advantages that published articles bring to the mechanisms of internalization and minimization of search costs is that the authors of journal articles describe their methods in detail. In fact, even the scientific instruments used (e.g. assays) are listed in many life science journals (Bergenholtz, 2014) to ensure reliability and reproducibility of findings. Accordingly, when the methods are presented in detail the amount of tacit knowledge surrounding a publication lessens and articles become a more relevant basis to transform internal tacit knowledge to codified knowledge, boost the exchange between tacit and codified knowledge, conduct new research, replicate research findings, or even avoid the costs of conducting duplicate research.

We therefore expect emerging life science firms, as well as other knowledge-based firms, to exploit the knowledge from academic research reported in journal articles to improve the quality of their innovation. Previous studies, not focusing on small firm research, have generally supported this expectation (Acosta et al., 2012; Sapsalis and van Pottelsberghe de la Potterie, 2007; Sorenson and Fleming, 2004).

Against this background, we form our first hypothesis:

Hypothesis 1: Firms that use academic science as an external input to their innovation process will produce innovations of higher quality.

Given the hypothesis that academic science and resulting knowledge *per se* can augment the quality of industrial innovation, the follow up question is why higher quality of academic science would improve the quality of industrial innovation. To address this question, we reflect upon the two main mechanisms by which we expect external academic knowledge to impact firm innovation: internalization and minimization of search costs. Under internalization, external academic science can assist in transforming in-house tacit knowledge to codified knowledge when it serves as an example and/or when it is used as an input to the knowledge transformation process. When examples and/or inputs of this kind are of higher quality, firm inventors are more likely to be confronted with more inspiring cases, to become exposed to better described, more efficient, more novel techniques and so on. Therefore, it is the increased usefulness, novelty and higher rigor facets of higher science quality that we expect to matter the most under the mechanism of internalization. Equivalently, other aspects of science quality such as higher societal impact and serving as foundation for follow-up academic works are less relevant. To illustrate, we would expect an emerging firm to gain more insights when using as external inputs the seminal articles of Stanley Cohen and Herber Boyer on DNA cloning and recombinant DNA (Cohen et al., 1972; Cohen et al., 1973) – both highly cited – when compared to lesser quality articles on similar topics as they describe more novel, and rigorous techniques that are useful for a wide array of applications. It is important to note that such mechanisms are more relevant for cases where academic science is used as an external knowledge source. When producing (academic) science in-house the mechanisms relate more to the internal generation of tacit knowledge (e.g. Hicks, 1995) and less so on the use of codified knowledge as a means to innovate.

Higher quality external academic science can also bring about significant benefits in search cost minimization. An important insight from conceptualizing science as a description of the technological landscape is that science can equip firm inventors with an understanding of the essential factors shaping the landscape (Fleming and Sorenson, 2004). When such understanding takes place, inventors can proceed in different ways than they would have proceeded in its absence; this process typically boosts innovation performance (Fleming and Sorenson, 2004). Starting then from the notion that higher science quality refers to work that is more groundbreaking, more rigorous and more insightful, we would expect this type of science to better describe the technological landscape as it will likely represent more novel ideas, techniques and scientific advancements. When this happens, inventors are equipped with a better understanding of their field, can follow more promising directions, and consequently produce innovations of higher quality. Again, then, similar to the mechanism of internalization it is the usefulness, rigor and novelty aspects of higher quality science that we expect to matter the most under cost minimization.

Turning our attention to empirical evidence, there is some support in the literature for the notion that knowledge heterogeneity, partly captured within quality differentials, may have a separate impact on innovation. One stream of studies documents that higher quality academics are generally more likely to produce inventions with higher commercial potential (Thursby et al., 2001) and are generally more apt to commercialize their innovations through various means, including firm creation (Zucker et al., 1998). Another stream of literature analyzes the spillover effects of academic and industrial R&D on local firm creation and reports that *who* conducts R&D matters for the rate of firm creation (Bade and Nerlinger, 2000; Karlsson and Nyström, 2011; Kolympiris et al., 2014). As well, studies that examine knowledge-based firms in regional clusters, indicate that the efficacy of knowledge flows in improving firm performance hinges upon the quality of the actors from whom the knowledge arises (Beaudry and Breschi, 2003; Kolympiris et al., 2011). All in all, it is conceivable that a similar effect is present in the case of knowledge flows from academic research to industrial innovation. We therefore hypothesize that:

Hypothesis 2: Firms that use higher quality academic science as an external input to their innovation process will produce innovations of higher quality.

The foregoing discussion suggests that academic science and particularly science of higher quality can lead to industrial innovation of higher quality. Still, there are contingencies under which such relationship holds.

A key insight of the work on a firm's absorptive capacity (Cohen and Levinthal, 1990) is that not all firms are equally equipped to benefit from external knowledge. Some firms are better able to accumulate and utilize external knowledge largely because they develop certain competencies that first allow them to understand it and then to embed it in their internal production function (Laursen and Salter, 2004). Indeed, for the case of academic science, Kealey and Ricketts (2014) build their conceptualization of science as a contribution good on the notion that the cognitive difficulty that some actors have in processing academic knowledge makes it partly excludable and hence a contribution good.[7]

But how can firms overcome this obstacle and develop higher cognitive abilities that can allow them to navigate the knowledge landscape map that science offers? One strategic competency that can allow firms in knowledge intensive industries to gain from academic research is the degree that the firm engages in basic research (Cockburn and Henderson, 1998; Fabrizio, 2009; Gambardella, 1992; Spencer, 2001). Given that a large part of published academic science is of basic nature, firms conducting in-house basic research are afforded several advantages (Rosenberg, 1990). For instance, they are better able to understand the assumptions underlying the published results, develop networks with academic scientists and thereby better identify the boundary conditions of those results, and ultimately better appreciate their applicability for in-house accumulation of knowledge (Lane and Lubatkin, 1998). Further,

[7] Kealey and Ricketts (2014) distinguish contribution goods from club goods by the nature of their excludability. Exclusion in contribution goods is based primarily on the capabilities of the prospective user rather than the property rights of the current possessor.

by conducting basic science in-house a given firm can become more familiar with academic research and share similar techniques, understandings and vocabulary with academia (Fabrizio, 2009); as a result knowledge transfer between academia and industry improves. Along the same lines, firms that promote academic publications by employees can be more effective in incorporating novel techniques in their toolbox and ultimately be more successful (Cockburn and Henderson, 1998; Henderson, 1994; Henderson and Cockburn, 1994). As Rosenberg (1990) puts it "A firm is much less likely to benefit from university research unless it also performs some basic research".

These sorts of advantages suggest that the type of research conducted in-house has a moderating effect on the relationship between academic research, its quality and the quality of industrial innovation. We therefore hypothesize that:

Hypothesis 3a: The degree that a given firm conducts more basic research in-house positively moderates the relationship between the use of external academic science by the firm and the quality of its innovations.

Hypothesis 3b: The degree that a given firm conducts more basic research in-house positively moderates the relationship between the quality of external academic science used by the firm and the quality of its innovations.

3. Methods and procedures

In order to empirically test the stated hypotheses we measure industrial innovations with patents and we relate, via econometric models, the quality of patents with measures of academic science intensity and quality while controlling for remaining factors that can also contribute to patent quality.

3.1. Dependent Variable

The dependent variable we employ in the analysis is the number of times a given patent is cited by later patents (forward citations). The intuition behind this measure is that higher citation levels imply superior significance or applicability. Indeed, a number of studies have consistently shown that forward citations correlate strongly with realized market value for a particular patent and are often regarded a reliable measure of patent quality (e.g. Fischer and Leidinger, 2014; Gambardella et al., 2008; Harhoff et al., 1999; Harhoff et al., 2003). Notably, forward citations also correlate with the size of the inventive step for a given technology: an additional citation correlates with meaningful increases in the innovativeness of a given technology (Moser et al., 2014).

Nevertheless, citation counts are not directly comparable across all patents; the number of citations is partly dependent on the age of the patent. More recent patents tend to have fewer citations largely due to the effective time needed before they become visible. In the same vein, the secular increase in the annual number of patents over time implies that very early patents may also tend to have fewer citations than more recent patents. Most patents receive their citations in the first few years after issue. If we look at a patent issued in, say, 1980, there were fewer patents of all classes issued in the succeeding five years than there were for a patent issued 25 years later, in 2005. All other things being equal, then, the earlier patent should have fewer forward citations simply because there were fewer other patents available to cite it (Lanjouw and Schankerman, 2004a). To address those two issues we use as our dependent variable the number of forward citations acquired by each sample patent as a proportion of the average number of citations received by all sample patents granted in the same year (CITECOUNT)[8]. Thus for any given year the average patent will have a value of 1 on this variable. A value of 2 indicates that the patent garnered twice the average number of forward citations, while a value of 0.5 would indicate it received only half the average number, and so forth.

3.2. Independent Variables

To test H1 we include in the analysis a variable that measures the number of journal articles that are referenced in the focal patent (JARTs). This choice is based on the findings that journal articles are a prime means via which industrial actors tap into knowledge from academic research and are a reliable proxy of knowledge flows from academia (Cohen et al., 2002; Roach and Cohen, 2013). The choice is also congruent with a number of previous studies that have used the number of scientific non-patent references as an indicator of the science intensity represented in a patent (Gittelman and Kogut, 2003; Roach and Cohen, 2013; Sapsalis and van Pottelsberghe de la Potterie, 2007; Sorenson and Fleming, 2004).

To test H2 we employ two novel variables. The first measures the average number of citations received by the articles referenced in each sample patent (ARTCITEs).[9] The second measures the impact factor of the journals in which these articles are published (IMPACT). The rationale for using two separate measures is twofold. First, as discussed before, knowledge

[8] We use the sample patents to calculate CITECOUNT because we expect the patents of the sample firms to be comparable and, thus, to have the same prospects in generating a comparable number of forward citations. We base such expectation on the largely homogeneous set of firms in terms of age, innovation potential and scientific field.
[9] We opt to use the total number of citations up to present and not citations per year to escape the following concern associated with citations per year: to illustrate, we use an example in which we employ citations from publication date to present as our measure of interest and assume that two articles receive 100 citations each. The first was published in 1990 and the second in 2000. If we employ citations per year the article published in 1990 will get a lower score than the one published in 2000 even though both received the same citations. As such, we would systematically undervalue older articles. Alternatively, we could calculate average citations from article publication date to patent grant year. Then, the article published in 1990 with 100 cites will score 100 when it is referenced by a patent from 1991, but only score 20 when referenced by a patent from 1995. As such, we would be assigning different scores to the same article.

quality is an encompassing term which almost by definition is difficult to capture with precision with a single measure. As such, by using two measures we aim to better describe knowledge quality. Second, while closely related, impact factors and citation rates may be capturing different aspects of quality: impact factors can better represent a form of pre-publication assessment of quality while citations a post-publication assessment (Eyre-Walker and Stoletzki, 2013). Accordingly, by using both measures we expect to account for such differences.

To be more precise, we use ARTCITEs under the premise that individual citation counts are superior to other measures as they are a direct measure of the quality of a specific journal article and thus are less prone to the statistical issues that may plague aggregated measures (Garfield, 2006). On the other hand, quality may not be always the main driver behind the number of citations for a given article; the academic preeminence of the author, network effects, exposure and other factors may also play a role. We expect IMPACT to cope with that issue. A journal's impact rating, as calculated by Journal Citation Reports (JCR), is generally accepted as a measure of the quality and importance of the *overall* research that a given journal publishes (Garfield, 2006). Indeed, partly because of the peer review process and partly because of the fact that tenure decisions, professional prestige, prizes, research grants and other forms of professional development in academia rely on articles published in journals with higher impact factors, what is typically reported in higher impact factor journals is the most noteworthy new findings and scholarly contributions. In practical terms, the fact that IMPACT approximates quality from a more aggregate perspective allows us to detect research quality even for articles that did not garner many citations but were published in journals hosting higher quality research.

To test H3a and H3b we interact the JARTs, ARTCITEs and IMPACT variables with the RESLEV variable that measures whether the firm conducts more applied or more basic science.

Following the methodology of Boyack et al. (2014), which is based on a textual analysis of the title and abstract of each article, we first calculate the probability that a given article published by the focal firm describes a) applied technology, b) engineering – technological mix, c) applied research and d) basic scientific research. Assigning a value of 1 for applied technology, 2 for engineering and so on in the next step we assign each article a research level score corresponding to the category with the highest probability. To construct RESLEV we average those values over all articles published by each firm. As such, increasing values of RESLEV correspond to firms conducting more basic research. For firms without any publications RESLEV takes the value of 0 under the assumption that firms that do not publish are not particularly active in conducting basic science. In line with H3a and H3b we expect positive coefficients for the interaction terms.

3.3. Control Variables

Guided by previous literature, we include a number of control variables that have been shown to impact patent quality. Reference to prior art in the form of patents is often considered an indicator of knowledge flows (Alcácer and Gittelman, 2006; Nemet and Johnson, 2012; Sampat and Ziedonis, 2004; Thompson, 2006). Accordingly, we include a variable that measures the number of prior art patent references for each patent (PATREFS) and expect a positive sign for the associated coefficient. To account for the quality of those patents we include a variable that measures the average number of total citations these patent have garnered over time (REFQUAL). Along the same lines, patents with broader scope of coverage are often regarded as more foundational and of higher quality (Czarnitzki et al., 2011; Lerner, 1994), although empirical results have been mixed (Harhoff et al., 2003). We measure scope as the number of distinct four-digit IPC categories assigned to a particular patent (SCOPE).

Patented innovations that are more original and groundbreaking in general also tend to be of higher quality. To measure originality we follow Harhoff and Wagner (2009) in constructing the originality measure first pioneered by Trajtenberg et al. (1997) (ORIGINAL). This index is a Herfindahl-type measure that measures the degree of similarity between the area of technology of the focal patent and the areas of technology of the patents referenced as prior art. More original patents tend to have a lesser degree of technological overlap with the patent they cite as, by definition, they represent more novel inventions. Formally, the measure is calculated as $ORIGINAL_i = 1 - \sum_{k=1}^{N_k} \left(\frac{Refs_{ik}}{Refs_i}\right)^2$, where patent *i* references patents from *k* technology classes. The originality index requires some adjustment from its basic form to enhance its accuracy. As in Trajtenberg et al. (1997) we adjust the final index for patents with more than two references by the factor $Refs_i/(Refs_i - 1)$. In addition, the index is undefined in the case of no references and always equals zero in the case of one reference. Arguably, patents with little or no prior art are among the most original innovations. Thus, in this study all patents with no prior art patent references are assigned an originality score of one. For patents with one reference, we assume that innovations whose prior art is in a different technology area as more original than one whose prior art is in the same area. Thus, for such cases we assign patents in the former category an originality score of one, while those in the latter receive a score of zero. More original patents also tend to occur earlier in the technology cycle (Régibeau and Rockett, 2010). To account for such observation, we include a measure of when the patent was issued (PATAGE), defined here as the number of months between the month the patent was granted and December 2012.

With regards to firm-specific features we include the number of inventors listed on each patent (INVENTS) as a proxy for firm resources devoted to a particular innovation; based on the premise that more resources tend to relate to higher quality we expect a positive sign for the corresponding coefficient. Other firm characteristics may have an impact on patent quality as well. In particular, younger firms may have a structure and culture that is more conducive to innovation than older firms that may be more bureaucratic (Acs et al., 1994), which then could lead to older firms producing innovations of lower quality (Balasubramanian and Lee, 2008). On the other hand, older organizations may be more innovative due to experience associated with a learning curve or because they are more able to capitalize on their research efforts. To account for such considerations we include the age of the firm in years in the year the patent was granted (FIRMAGE).

We include three additional firm-specific characteristics. The first counts the number of articles published by firm scientists prior to the application year of each patent in our database (PUBCOUNT). We expect firms whose employees are engaged in academic research to be able to understand external academic research more comprehensively and as a result to utilize this external knowledge to produce innovations of higher quality. As argued in the discussion leading to hypothesis 2 research quality is not homogeneous and quality differentials can lead to different innovation outcomes. As such, the second firm-specific characteristic we include in the analysis is the average number of citations received by articles authored by firm employees (AVGCITES). The third firm-specific variable accounts for the increased scope of collaborations between universities and industrial actors (Link and Rees, 1990). Given such scope we need to account for the degree of intentional knowledge transfer between the sample firms and academic institutions. To do so, we include a binary variable that reflects whether a given patent includes a university as a co-assignee (COLLAB). COLLAB also accounts for the possibility that planned knowledge transfer can be more effective than journal articles in transmitting tacit knowledge. Theoretically, because university research tends to be more basic it can result in more valuable innovations (Czarnitzki et al., 2011). However, firms that engage in publishing can suffer in terms of innovative output (Gittelman and Kogut, 2003). Thus, university collaboration can steer the quality of firm innovation either way, so it is difficult to anticipate the sign of COLLAB *a priori*. To account for any unobservable characteristics of a given firm such as organizational culture or innate ability of employees that can also drive innovation quality we include firm-specific dummy variables in the analysis. Finally, to address differences in citation patterns across technology areas, we include a set of binary variables that represent the technology class to which the focal patent belongs.

4. Data sources and presentation

4.1. Identifying Emerging Firms through the Small Business Innovation Research Program

We identified emerging life sciences firms via grant recipients of the Small Business Innovation Research (SBIR) program. In recognition of the role small- and medium-sized enterprises (SMEs) play in innovation, the US Congress established the SBIR program in 1982 (Audretsch, 2002; Cooper, 2003). The legislation mandates a percentage of federal extramural research and development expenditures be allocated to small business firms. Funding is provided in two phases, followed by commercialization; further funding, if any, comes from private investors. In addition to supporting R&D in SMEs, the program has the stated goals of stimulating technological innovation in general and increasing the commercialization of innovations resulting from federal R&D programs. Pursuant to these goals, the bulk of SBIR funds are awarded to emerging firms operating in cutting edge knowledge-intensive research areas such as the life sciences, electronics, materials and energy conversion. As well, the majority of the recipients are still at the (very) early stages of firm growth and as such the SBIR funds are typically the initial financial proceeds from external sources. Given these features of the SBIR program, it appears to be a suitable source of data for the present study; award winners are typically emerging, innovative, independent firms with limited financial resources (Audretsch, 2003; Audretsch et al., 2002).[10] As such, seeking academic knowledge via a spillover process appears likely for this cohort of firms.

4.2. Data description

To develop our data set we began with the entire population of the 1671 life science firms that received an SBIR award from 1983 to 2006, as provided by InKnowVation. We searched for award recipients up to 2006 (and applications submitted up to 2007) in order to be able to observe and source the full list of patents granted to those firms and minimize potential truncation of the data. Given the substantial time lag between the application and the grant date of a given patent (Mitra-Kahn et al., 2013), using 2007 as the cut-off point of our applications allowed us to study patents granted through the end of 2012.

We then searched the patent database maintained by commercial vendor Thompson Innovation[11] to identify which of the 1671 SBIR award winners had filed patent applications from 1971 through 2007. This procedure reduced the sample to the 910 firms that had patent activity for the time period in question[12]. These 910 firms had been granted 15,505 patents. These 15,505 patents cited more than 575,000 non-patent references which included articles in academic journals, court documents, industry reports, laboratory reference books, gene bank entries and others. The inconsistent format of non-patent references in the patent files prompted

[10] Note that by concentrating on a relatively homogeneous set of firms allows us the opportunity to eliminate sources of patent heterogeneity arising from firm and industry characteristics, such as patent portfolio size (Lanjouw and Schankerman, 2004b) and market structure (Ziedonis, 2004) as well as to lessen concerns that unobserved heterogeneity can hamper our estimates in a meaningful way.

[11] To ensure we measure all relevant patents and to cope with potential issues that could arise from differences in the way a given firm is reported as the assignee in different patents we tried a number of variants in the firm name (e.g. 20/20 Genesystems was coded as 20/20 Genesystems Inc., 20/20 Gene*, 20-20 Genesystems and so on).

[12] Firms with patents were slightly larger and older than firms without patents and in general had sourced more funds from the SBIR program.

us to use a combination of manual cleanup and automated procedures to separate the journal articles from the rest of the non-patent references.[13] Manual cleanup and automated procedures both have potential for error, albeit of different kinds. Thus, to keep the data process tractable without loss of generality, and given that the unit of analysis is the patent, we randomly selected 7753 patents (half of the population of 15,505 patents) which we manually processed. These 7753 patents represented approximately 87 percent of the patents owned by 434 firms. We then manually identified and added the remaining 1156 patents for those firms to reach the final sample of 434 firms with 8908 patents.[14]

The Thomson Innovation dataset contained the information we needed to construct the dependent and independent variables (i.e. number of forward citations, list of patent and non-patent references, etc.). Once we isolated the journal articles from the full list of non-patent references we developed a list of journal titles, tied to each patent, which were then matched to a list of 8594 journal titles with International Standard Serial Numbers (ISSNs), from the Journal Citation Reports (JCR). Using the ISSNs we located the journals of interest, the articles included in the patent references and the number of citations per article from SCOPUS. More specifically, we exploited the option in SCOPUS to look up articles based on the affiliation of the authors; we

[13] More specifically, we faced two major obstacles. First, there was no standard order to the elements of the reference and no standard punctuation scheme. That is, journal articles are not reported in the same way at the USPTO not only across patents but also within the same patent, and hence in our data source. For instance, sometimes the title was placed between brackets, other times it was placed between quotation marks/periods/commas and other times it was not placed between any kind of identifiers. Second, across and within patents, journal titles are not reported consistently. To illustrate, we came across 311(!) different ways of abbreviating the journal title "Proceedings of the National Academy of Sciences of the USA". The most common punctuation practice was to place the article title in quotes, so we used this criteria to separate each reference into three segments – before, within, and after quotes. We then used an automated search procedure to initially identify journal titles, followed by manual cleanup. We identified journal titles for 254,827 references, over 90% of those initially identified as journal articles.

[14] Firms included in this sample were largely similar to those left out. The main exception was that the firms in the sample had more patents per firm on average (19.3 vs. 10.3). The difference is expected as the firms with the more patents are more heavily represented in the pool of patents we randomly draw from (precisely because they have more patents). As such, they are more likely to be more heavily represented in the set of firms we study.

input the firm names (as well as variants of the name to make sure we did not miss any articles) in the affiliation search and then downloaded all the articles in which at least one author was affiliated with the focal firm.

To calculate the ORIGINAL index we first converted the primary IPC code of the focal patent and all of its referenced patents to the 35 category ISI-OST-INPI classification, a classification scheme that more accurately reflects technological relatedness than does the IPC coding system (Schmoch, 2008). Each term in the summation, then, is the number of patent references belonging to a particular technology class divided by the total number of patent references for the focal patent.[15] The index ranges from 0 for the least original patents to 1 for the most original patents. We also used the classification scheme which assigned IPCs to technology areas to construct the technology class dummy variables.

Table 1 displays the descriptive statistics for the variables used in the analysis. As described above, the average patent in any given grant year has a value of one for the dependent variable, CITECOUNT. Thus, by construction this variable has a mean of 1. CITECOUNT has a somewhat skewed distribution, as the median is roughly one-third the mean and the modal value is zero[16]. The heterogeneity of the observed values implies that our sample is composed of patents of different quality.

< TABLE 1 ABOUT HERE>

With regards to the variables we employ to test our hypotheses, the average patent in our data set had 25 journal articles in its list of references, with half of the patents having fewer than 11 journal articles in their reference list. In general the 25 figure we refer to above strengthens our

---

[15] In this sample a total of 549 unique four digit IPC codes were collapsed into 33 technology classes. Two technology classes were not represented in this data set.

[16] Note that to calculate the variable we employed the number of forward citations of the patents by the original 910 firms. Even when the variable is constructed by employing the patents of the 434 sample firms, the results remain qualitatively similar.

expectation that emerging life sciences firms rely on academic science as a source of knowledge and it is in line with the note of McMillan et al. (2000) that biotechnology patents reference basic research more heavily than other areas of technology. The cited articles were published on journals with an average impact factor of 9.42 with a minimum of 0.23 and a maximum of 50.81. As well, the modal value of the ARTCITEs variable suggests that most articles referenced in the sample patents had not received any citations even though there is substantial variability (the standard deviation is 2.5 times larger than the mean value). The wide distribution of the variables we use to approximate science quality is important as it indicates the heterogeneity of the quality of the science focal firms use as an external input to their innovative process.[17]

The average sample patent cited 37 patents, belonged to 3 IPC categories, had an originality score of 0.61 (out of 1), and had 4 inventors. Information not reported in Table 1 illustrates the young age of the sample firms: half of them were 2 years old when they received the SBIR award while the modal value for firm age at the receipt of the award was 1 year.[18]

With regards to the publication activity of the sample firms, most firms had not published any articles. However, three firms had well above average publication rates (more than 830 publications each) and their record in essence inflates the reported average value of the variable in question. As shown in Appendix Table 2 where we omit these firms from the analysis we observe only small differences when compared to the baseline estimates presented in Table 3. On average the articles published by the sample firms had garnered 46 citations. Interestingly, half of the articles published by the sample firms appeared in journals scoring high in the level of

[17] These statistics alleviate concerns that our sample is subject to the general finding that higher quality science tends to appear more frequently as prior art in patent documents (Narin et al., 1997). Along the same lines, even if such possibility holds in our sample, it would imply that the coefficients of IMPACT and ARTCITEs represent a *lower* bound of the true effect of science quality on the quality of industrial innovation.

[18] Note that the statistics of firm age at patent application as reported in Table 1 are somewhat inflated mostly because of a handful of older and larger firms that had a large number of applications. As such, the patents of those firms are represented in the sample more heavily.

basicness (3.24 out 4). This statistic reinforces the close ties between basic research and applied work within the life sciences.

The correlation coefficients shown in Table 2 are, for the most part, relatively small, which should help us to estimate the separate effects of each right-hand variable on patent quality.

< TABLE 2 ABOUT HERE>

Before we proceed to the empirical results, we need to discuss two important considerations that relate to the suitability of the sample at hand and the interpretation of the results.

The first consideration is that the observations from the 434 firms we employ in the analysis are relatively similar to the observations from the 910 firms that composed the population of life sciences firms that received SBIR grants and patented their research. Appendix Table 1 presents the descriptive statistics for the variables we employ in the empirical specifications (except IMPACT, JARTs, and selected firm-specific variables) for the 910 firms, and Table 1 presents the descriptive statistics for the 434 firms. The variables in these two tables exhibit only small differences, leading us to conclude that the sample of firms in this study is representative of the original group.

The second consideration relates to the degree that our estimates reflect knowledge spillovers from academia to industry. In total, employees of our sample firms were listed as (co)-authors of 19,479 articles listed in SCOPUS. As previously mentioned, our sample patents cited 254,827 articles. As such, if all 19,479 articles are cited in the focal patents, only 7.6% of the cited articles represent academic science conducted within the firm holding the patent, hence not a spillover process. Therefore, for the full sample we estimate that at most 8 out of 100 articles do not capture knowledge spillovers. However, this is an upper bound and does not necessarily hold for every patent in the analysis. Our estimates could be subject to noise if there is significant variability among patents in the percentage of cited articles that actually represent knowledge spillovers. Because of the previously described non-standard format that JARTs appear in the patent files we are not able to precisely estimate the relevant percentage. However, to assess the issue we manually looked up the reference lists of more than 600 randomly selected patents; no patent had more than 10 percent of its cited articles co-authored by the assignee firm and for most patents the equivalent percentage was well below the 7.6 figure discussed above. These low percentages are significant for the interpretation of the JARTs, IMPACT and ARTCITE coefficients in that they strongly imply that for the largest part these coefficients do in fact measure knowledge flows from academia to industry rather than internally produced academic research. Relatedly, we maintain that our estimates reflect primarily knowledge spillovers even when considering that the prior art listed in patents is added both by the applicant firm and by the patent examiner. While examiners do influence the list of cited literature in a given patent (Alcácer and Gittelman, 2006; Alcácer et al., 2009), they do so only to a minimal degree when it comes to non-patent references (Tijssen, 2002) perhaps because they are less familiar with the academic literature when compared to their exposure to patents.[19] Specifically for life sciences, and without separating patent and non-patent references, evidence shows that life sciences is a field where the share of cited references that come directly from the inventors is higher than the corresponding share in the majority of other fields (Alcácer et al., 2009; Criscuolo and Verspagen, 2008). All in all, such observations are among the prime drivers in suggesting that journal articles cited in a given patent are a reliable measure of knowledge flows (Roach and Cohen, 2013).

[19] A substantial portion of the sample patents were granted before 2001, which is the year after which USPTO makes available information about whether a cited reference was added by the examiner or the applicant. Accordingly, we cannot check who added the journal article in the list of patent references.

5. Empirical Results

5.1. Main Results

Table 3 presents the Tobit estimates where the dependent variable is the adjusted number of forward citations for a given patent. Because, as a proportion, this variable has a lower bound of 0 we use a Tobit estimator to address this feature of the data (Wooldridge, 2009). Model 1 includes only the variables used to test Hypotheses 1 and 2. In model 2 we add the patent-specific control variables to model 1 and in model 3 we add the firm-specific control variables. Model 4 is the fully specified model without testing H3 (i.e. without the moderating effects). In models 5 and 6 we test H3 and in model 7 we test H3a. Model 8 is the fully specified model testing all hypotheses.

The multicollinearity condition index and the fit statistics reported for all models suggest that a) inference concerns associated with multicollinearity are not pronounced since across specifications the index is within the range of the threshold level of 30 (Belsley et al., 1980) and b) that the models have explanatory power.

< TABLE 3 ABOUT HERE>

Across specifications we fail to reject H1. JARTs shows a positive and statistically significant coefficient; the higher the science intensity of the firm patents the higher their quality. As discussed previously, the fact that the vast majority of the cited articles were not co-authored by the sample firms indicates that the reported coefficients measure knowledge spillovers from academia to industry.

We find the results pertaining to H2 and H3 particularly interesting. The findings of model 5 support H3b. Firms that conduct more basic research are better able to use higher quality academic science, as measured by the IMPACT[20] variable, to produce innovations of higher quality. In fact, the basicness of firm research has a strong moderating effect on the relationship between higher quality academic science and higher quality industrial innovation: the marginal effect of IMPACT on the quality of industrial innovation is insignificant when RESLEV takes on the value of 0 (models 5 and 8); that is when the firm does not conduct basic research. All in all, we find that the separate effects of higher quality academic science accrue only to firms that can actually better accumulate such science via their basic science orientation.

We do not find a similar effect for science intensity as the interaction term in model 7 is statistically insignificant (hence we reject H3a). An explanation in line with these findings is that higher quality academic science may be more difficult to grasp than lesser quality science and as such conducting basic science may afford additional benefits only for the case of higher quality academic research.

On aggregate, these results demonstrate that science intensity has an independent positive effect on the quality of patented innovations and so does science quality, at least for the cohort of firms that are active in conducting more basic science. This latter result is important in that it highlights that cognitive difficulties and absorptive capacity can determine which types of firms can benefit from academic science of higher quality via a knowledge spillover process. Upon reflection, these results give support to our hypotheses and are in line with our discussion that science quality needs to be at the center of studies analyzing the links between academic research and industrial innovation.

[20] The results of model 6 do not reveal similar findings for the ARTCITEs variable. A possible driver behind that finding is that individual citations may be driven by factors other than quality; because, as we explain in section 3.2, the impact factor is less prone to such issues, it may be capturing knowledge quality in a more aggregate manner.

As it pertains to the control variables, patents that cite more patents as prior art are of higher quality, while patent scope and originality have no discernable effect. The positive sign of the variable measuring the quality of the referenced patents (REFQUAL) reinforces our main conclusion that higher quality inputs lead to higher quality innovations. Firm-specific variables showed varying influence on quality. In line with previous works (Czarnitzki et al., 2011), the number of inventors listed on a patent, representing the firm's resources devoted to R&D, showed a significant positive relationship with patent quality. Along the same lines, younger firms had higher quality patents as well. The negative sign of the COLLAB variable implies that, for the sample firms, collaboration with universities did not improve but rather degraded the quality of innovations. Firms publishing better, rather than more, articles also had patents of higher quality.

5.2. Sensitivity Analysis

In Table 4 we present models that test the robustness of our findings.

< TABLE 4 ABOUT HERE>

As we explain in footnote 9 we opt to use the total number of citations up to present as one of our measures of article quality and not citations per year so that we do not systematically undervalue older articles. In model 1 we test the sensitivity of our conclusions to the construction of the variable in question. We replace ARTCITEs in model 8 of Table 3 with a variable that measures citations of a focal article as a proportion of the average number of citations received by all sample articles published in the same year. Thus, for any given year the average article will have a value of 1 on this variable. A value of 2 indicates that the article garnered twice the average number of citations, while a value of 0.5 would indicate it received only half the average number, and so forth. The results remain qualitatively similar to the baseline estimates.

If at a certain point in time firm scientists explore a pertinent technology of explorative nature and if such technology requires them to draw on high quality science, then such time-variant unobserved characteristics of the opportunities being explored can determine both the quality of science being used as an input and the quality of the patents (the output). In such case, our findings would not be reflecting a link between the quality of science and the quality of industrial innovation but would be a result of the unobserved heterogeneity described above. We expect the technological class dummy variables to account for the nature of the technology being explored. To check whether the temporal dimension of the project a firm is exploring influences our estimates, in model 2 of Table 4 we include a set of dummy variables that match the application year of the patent[21] and in model 3 we replace the technological class dummy with the year dummies. We do not observe significant differences between these models and the baseline estimates we present in table 3.

In model 4 of Table 4 we drop the two science quality measures from the analysis but keep JARTs. We do so in order to further demonstrate that explicitly accounting for science quality is a meaningful exercise in explaining the quality of firm innovation. That is, if the coefficient of JARTs would pick up the quality effects when the quality measures are not included in the analysis we would expect it to increase in magnitude substantially. We do not observe such change as the magnitude of the JARTs coefficient remains unchanged (0.004). This consistency illustrates that incorporating knowledge quality in the analysis helps in better understanding the quality of firm innovation.

[21] The results remain unchanged even if we define the year dummies at t-1 or t-2 with t being the application year.

6. Conclusions and discussion

In the introduction of the paper we noted that the contribution of knowledge from academic research on industrial innovation at the firm level has been examined in various studies but the impact of the heterogeneity of such knowledge has not been assessed. Further, in this body of literature the contributions of academic knowledge on industrial innovation via a knowledge spillover process has not received attention in any significant depth. For these reasons, in this study we examine whether different quality knowledge from academic research leads to industrial innovation of different quality via knowledge spillovers. Importantly, we also uncover the boundaries that relate to firm-specific characteristics and allow a given firm to more effectively improve its innovation record using academic science of higher quality as an input.

Empirically, we use a micro level analysis and we show that patents of entrepreneurial life science firms that reference academic research published in journal articles are of higher quality. We also document a separate significant effect of knowledge quality. We find that articles published in journals with a higher impact factor lead to industrial innovation of higher quality as long as they are employed by firms active in conducting and publishing basic research. Our results, therefore, suggest that better equipped firms can more effectively exploit the insights provided by higher quality academic science. On the other hand, we do not document differences across different cohorts of firms in the exploitation of academic science *per se*. As such, the case might be made that higher quality academic science is more challenging to grasp and the higher cognitive ability afforded to firms with a basic science orientation gives them an advantage in that respect.

From a more aggregate perspective, we complement a long stream of research highlighting the beneficial effects of developing higher absorptive capacity; our estimates imply that increased cognitive capabilities achieved via in-house basic research allow firms to gain from academic science of higher quality. As such, follow up studies can further investigate the links between academic research quality and industrial innovation. Our results also inform the conceptualization of science as a contribution good (Kealey and Ricketts, 2014). This conceptualization challenges the non-excludable nature of academic science as it emphasizes that cognitive difficulties in processing knowledge can act as a barrier to access. Simply put, to use the science, one must be able to understand it. Our estimates paint a more nuanced story: not science *per se* but higher quality, and perhaps more difficult to grasp, science appears to exhibit contribution good characteristics. We see this as a fruitful area for further research.

Our study contributes to the literature that examines how academic science impacts industrial innovation. In the face of tighter fiscal policies and generally shrinking academic research budgets universities are increasingly expected to generate commercial outcomes from on-campus research. Such expectations often translate to more demands on faculty to exploit their research in the market. However, commercial demands often come on top of pressures to publish in high quality journals as tenure decisions, professional progress and the like are closely linked to publication outcomes. Our results imply that the dual demands of higher quality publications and commercial exploitation can be reconciled. Publication outcomes of higher quality, i.e. those correlating strongly with career advancement, are also the ones that can lead to higher quality industrial innovation. Such finding can, therefore, inform university administrators, public granting agencies promoting innovation, directors of development policies and directors of technology transfer at different universities when designing frameworks and policies relating to university based innovation, tenure decisions and the like. Similarly, the fact that emerging young entrepreneurial firms are among the prime drivers of economic growth implies that our study speaks directly to all policy makers that envision boosting economies by encouraging small firm innovation.

Our work also suggests that efforts by government officials and researchers to standardize the process of reporting, identifying and utilizing scientific non-patent references could be a worthwhile exercise (Callaert et al., 2012). References listed in patent documents represent a rich source of information for studies on innovation, knowledge exchange and the like but the non-standard format in which they are reported can hamper their usefulness and discourage researchers to address in depth the linkages between academic knowledge quality and commercial outcomes.

In terms of managerial implications, we reinforce the conclusion reached in a large part of the literature that investing in basic research may pay off for emerging knowledge intensive firms. Previous works highlight that in-house basic research assists firms in many ways including the exploitation of academic research. Our work enriches this finding. We bring to light the novel finding that in-house basic research is instrumental in exploiting higher quality academic research; such research can lead to higher quality firm innovation but is likely more difficult to comprehend and use. This is a point worth emphasizing when considering that the resource scarcity entrepreneurial firms often face may discourage them from investing in basic research despite their high intellectual capital. While our estimates are not meant to provide a cost benefit analysis regarding the increases in innovation quality and the costs of conducting basic science, they do indicate that in-house basic science can afford significant benefits.

To close we acknowledge that our study is subject to limitations. For our work we have focused on emerging life sciences firms in part because of the welfare enhancing character of these firms but also because they present a suitable template for the research question at hand. As such, our results must be generalized to other sectors and different types of firms through further studies. Similarly, due to computational difficulties we were not able to precisely estimate how many of the cited journal articles are co-authored by the sample firms. Perhaps future studies can discover how to incorporate such considerations. More broadly, with an eye on the evidence demonstrating their suitability as the best available metric, we rely on assessments of research quality derived by subjective assessments of quality (e.g. how many authors have cited a piece of academic work which indirectly suggests their quality assessment). We are aware that the measures we employ are the best available but are still subject to shortcomings that make them far from a perfect measure (Eyre-Walker and Stoletzki, 2013). We are not, however, aware of measures that could more objectively and more comprehensively describe research quality for studies, like ours, that rely on large number of publications. As such, we encourage future studies to address the issue in depth.

References

Acosta, M., Coronado, D., Martínez, M., 2012. Spatial differences in the quality of university patenting: Do regions matter? Research Policy 41, 692-703.

Acs, Z.J., Anselin, L., Varga, A., 2002. Patents and innovation counts as measures of regional production of new knowledge. Research Policy 31, 1069-1085.

Acs, Z.J., Audretsch, D.B., Feldman, M.P., 1994. R&D spillovers and innovative activity. Managerial and Decision Economics 15, 131-138.

Aksnes, D.W., 2003. Characteristics of highly cited papers. Research Evaluation 12, 159-170.

Aksnes, D.W., 2006. Citation rates and perceptions of scientific contribution. Journal of the American Society for Information Science and Technology 57, 169-185.

Aksnes, D.W., Rip, A., 2009. Researchers' perceptions of citations. Research Policy 38, 895-905.

Alcácer, J., Gittelman, M., 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. Review of Economics and Statistics 88, 774-779.

Alcácer, J., Gittelman, M., Sampat, B., 2009. Applicant and examiner citations in U.S. patents: An overview and analysis. Research Policy 38, 415-427.

Almeida, P., Hohberger, J., Parada, P., 2011. Individual scientific collaborations and firm-level innovation. Industrial and Corporate Change 20, 1571-1599.

Argyres, N.S., Liebeskind, J.P., 1998. Privatizing the intellecutal commons: Universities and the commercialization of biotechnology. Journal of Economic Behavior & Organization 35, 427.

Arundel, A., Kabla, I., 1998. What percentage of innovations are patented? Empirical estimates for European firms. Research Policy 27, 127-141.

Audretsch, D.B., 2002. The dynamic role of small firms: Evidence from the US. Small Business Economics 18, 13-40.

Audretsch, D.B., 2003. Standing on the shoulders of midgets: The US Small Business Innovation Research program (SBIR). Small Business Economics 20, 129-135.

Audretsch, D.B., Link, A.N., Scott, J.T., 2002. Public/private technology partnerships: evaluating SBIR-supported research. Research Policy 31, 145-158.

Azoulay, P., Stuart, T., Wang, Y., 2014. Matthew: Effect or fable? Management Science 60, 92-109.

Bade, F.-J., Nerlinger, E.A., 2000. The Spatial Distribution of New Technology-Based Firms: Empirical Results for West-Germany. Papers in Regional Science 79, 155-176.

Balasubramanian, N., Lee, J., 2008. Firm age and innovation. Industrial and Corporate Change 17, 1019-1047.

Balconi, M., Brusoni, S., Orsenigo, L., 2010. In defence of the linear model: An essay. Research Policy 39, 1-13.

Beaudry, C., Breschi, S., 2003. Are firms in clusters really more innovative? Economics of Innovation and New Technology 12, 325-342.

Belenzon, S., Patacconi, A., 2014. How does Firm Size Moderate Firms' Ability to Benefit from Invention? Evidence from Patents and Scientific Publications. European Management Review 11, 21-45.

Belsley, D.A., Kuh, E., Welsch, R.E., 1980. Regression diagnostics. Wiley Online Library.

Bergenholtz, C., 2014. Uncovering a scientific practice: The role of references to scientific instrument firms in the natural sciences. European Management Review Forthcoming.

Boyack, K.W., Patek, M., Ungar, L.H., Yoon, P., Klavans, R., 2014. Classification of individual articles from all of science by research level. Journal of Informetrics 8, 1-12.

Callaert, J., Grouwels, J., Van Looy, B., 2012. Delineating the scientific footprint in technology: Identifying scientific publications within non-patent references. Scientometrics 91, 383-398.

Cassiman, B., Di Guardo, M.C., Valentini, G., 2010. Organizing links with science: Cooperate or contract?: A project-level analysis. Research Policy 39, 882-892.

Cassiman, B., Veugelers, R., Zuniga, P., 2008. In search of performance effects of (in) direct industry science links. Industrial and Corporate Change 17, 611-646.

Cockburn, I.M., Henderson, R.M., 1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. The Journal of Industrial Economics 46, 157-182.
Cohen, S.N., Chang, A.C., Hsu, L., 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. Proceedings of the National Academy of Sciences of the United States of America 69, 2110-2114.
Cohen, S.N., Chang, A.C.Y., Boyer, H.W., Helling, R.B., 1973. Construction of biologically functional bacterial plasmids in vitro. Proceedings of the National Academy of Sciences of the United States of America 70, 3240-3244.
Cohen, W.M., Levinthal, D.A., 1990. Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly, 128-152.
Cohen, W.M., Nelson, R.R., Walsh, J.P., 2002. Links and impacts: the influence of public research on industrial R&D. Management Science 48, 1-23.
Cooper, R.S., 2003. Purpose and performance of the small business innovation Research (SBIR) program. Small Business Economics 20, 137-151.
Criscuolo, P., Verspagen, B., 2008. Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. Research Policy 37, 1892-1908.
Cumming, D., Johan, S., Zhang, M., 2014. The Economic Impact of Entrepreneurship: Comparing International Datasets. 22, 162-178.
Czarnitzki, D., Hussinger, K., Schneider, C., 2011. Commercializing academic research: the quality of faculty patenting. Industrial and Corporate Change 20, 1403-1437.
de la Potterie, B.v.P., 2011. The quality factor in patent systems. Industrial and Corporate Change 20, 1755-1793.
Eyre-Walker, A., Stoletzki, N., 2013. The assessment of science: the relative merits of post-publication review, the impact factor, and the number of citations. PLoS biology 11, e1001675.
Fabrizio, K.R., 2009. Absorptive capacity and the search for innovation. Research Policy 38, 255-267.
Fischer, T., Leidinger, J., 2014. Testing patent value indicators on directly observed patent value—An empirical analysis of Ocean Tomo patent auctions. Research Policy 43, 519-529.
Fleming, L., Sorenson, O., 2004. Science as a map in technological search. Strategic Management Journal 25, 909-928.
Gambardella, A., 1992. Competitive advantages from in-house scientific research: the US pharmaceutical industry in the 1980s. Research Policy 21, 391-407.
Gambardella, A., Harhoff, D., Verspagen, B., 2008. The value of European patents. European Management Review 5, 69-84.
Garfield, E., 2006. The history and meaning of the journal impact factor. Jama 295, 90-93.
Gittelman, M., 2007. Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology. Organization Science 18, 724-741.
Gittelman, M., Kogut, B., 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. Management Science 49, 366-382.
Gray, P.H., Meister, D.B., 2004. Knowledge sourcing effectiveness. Management Science 50, 821-834.
Hagedoorn, J., Cloodt, M., 2003. Measuring innovative performance: is there an advantage in using multiple indicators? Research Policy 32, 1365-1379.
Haltiwanger, J., Jarmin, R.S., Miranda, J., 2013. Who creates jobs? Small versus large versus young. Review of Economics and Statistics 95, 347-361.
Harhoff, D., Narin, F., Scherer, F.M., Vopel, K., 1999. Citation frequency and the value of patented inventions. Review of Economics and Statistics 81, 511-515.
Harhoff, D., Scherer, F.M., Vopel, K., 2003. Citations, family size, opposition and the value of patent rights. Research Policy 32, 1343-1363.
Harhoff, D., Wagner, S., 2009. The duration of patent examination at the European Patent Office. Management Science 55, 1969-1984.
Henderson, R., 1994. The evolution of integrative capability: innovation in cardiovascular drug discovery. Industrial and Corporate Change 3, 607-630.

Henderson, R., Cockburn, I., 1994. Measuring competence? Exploring firm effects in pharmaceutical research. Strategic Management Journal 15, 63-84.
Henrekson, M., Johansson, D., 2010. Gazelles as job creators: a survey and interpretation of the evidence. Small Business Economics 35, 227-244.
Hicks, D., 1995. Published papers, tacit competencies and corporate management of the public/private character of knowledge. Industrial and Corporate Change 4, 401-424.
Igami, M., 2013. Patent Statistics as Innovation Indicators? Hard Evidence.
Jensen, R.A., Thursby, J.G., Thursby, M.C., 2003. Disclosure and licensing of University inventions:'The best we can do with the s** t we get to work with'. International journal of industrial organization 21, 1271-1300.
Jong, S., Slavova, K., 2014. When publications lead to products: The open science conundrum in new product development. Research Policy.
Karlsson, C., Nyström, K., 2011. Knowledge accessibility and new firm formation, in: Desai, S., Nijkamp, P., Stough, R.R. (Eds.), NEW DIRECTIONS IN REGIONAL ECONOMIC DEVELOPMENT. Edward Elgar Publishing, Inc., Northampton, Massachusetts, pp. 174-197.
Kealey, T., Ricketts, M., 2014. Modelling science as a contribution good. Research Policy 43, 1014-1024.
Khilji, S.E., Mroczkowski, T., Bernstein, B., 2006. From Invention to Innovation: Toward Developing an Integrated Innovation Model for Biotech Firms. Journal of product innovation management 23, 528-540.
Kolympiris, C., Kalaitzandonakes, N., 2013. Geographic scope of proximity effects among small life sciences firms. Small Business Economics 40, 1059-1086.
Kolympiris, C., Kalaitzandonakes, N., Miller, D., 2011. Spatial collocation and venture capital in the US biotechnology industry. Research Policy 40, 1188-1199.
Kolympiris, C., Kalaitzandonakes, N., Miller, D., 2014. Public funds and local biotechnology firm creation. Research Policy 43, 121-137.
Lane, P.J., Lubatkin, M., 1998. Relative absorptive capacity and interorganizational learning. Strategic Management Journal 19, 461-477.
Lanjouw, J.O., Schankerman, M., 1999. The quality of ideas: measuring innovation with multiple indicators. National bureau of economic research.
Lanjouw, J.O., Schankerman, M., 2004a. Patent quality and research productivity: Measuring innovation with multiple indicators. The Economic Journal 114, 441-465.
Lanjouw, J.O., Schankerman, M., 2004b. Protecting Intellectual Property Rights: Are Small Firms Handicapped? Journal of Law and Economics 47, 45-74.
Laursen, K., Salter, A., 2004. Searching high and low: what types of firms use universities as a source of innovation? Research Policy 33, 1201-1215.
Lawless, M., 2014. Age or size? Contributions to job creation. Small Business Economics 42, 815-830.
Leonard, D., Sensiper, S., 1998. The role of tacit knowledge in group innovation. California Management Review, 112-132.
Lerner, J., 1994. The importance of patent scope: an empirical analysis. RAND Journal of Economics 25, 319-333.
Li, D., Agha, L., 2015. Big names or big ideas: Do peer-review panels select the best science proposals? Science 348, 434-438.
Liebeskind, J.P., Oliver, A.L., Zucker, L., Brewer, M., 1996. Social networks, learning, and flexibility: Sourcing scientific knowledge in new biotechnology firms. Organization Science 7, 428-443.
Link, A.N., Rees, J., 1990. Firm size, university based research, and the returns to R&D. Small Business Economics 2, 25-31.
McMillan, G.S., Narin, F., Deeds, D.L., 2000. An analysis of the critical role of public science in innovation: the case of biotechnology. Research Policy 29, 1-8.
Mitra-Kahn, B., Marco, A., Carley, M., D'Agostino, P., Evans, P., Frey, C., Sultan, N., 2013. Patent backlogs, inventories and pendency: An international framework. IPO & USPTO joint report, http://www.ipo.gov.uk/pro-ipresearch.htm.

Moser, P., 2005. How do patent laws influence innovation? Evidence from nineteenth-century world's fairs. American Economic Review 95, 1214-1236.
Moser, P., Ohmstedt, J., Rhode, P.W., 2014. Patent Citations and the Size of the Inventive Step-Evidence from Hybrid Corn. Available at SSRN 1888191.
Nagaoka, S., 2007. Assessing the R&D management of a firm in terms of speed and science linkage: evidence from the US patents. Journal of Economics & Management Strategy 16, 129-156.
Narin, F., Hamilton, K.S., Olivastro, D., 1997. The increasing linkage between US technology and public science. Research Policy 26, 317-330.
Nemet, G.F., Johnson, E., 2012. Do important inventions benefit from knowledge originating in other technological domains? Research Policy 41, 190-200.
Neumark, D., Wall, B., Zhang, J., 2011. Do small businesses create more jobs? New evidence for the United States from the National Establishment Time Series. Review of Economics and Statistics 93, 16-29.
Nightingale, P., Coad, A., 2014. Muppets and gazelles: political and methodological biases in entrepreneurship research. Industrial and Corporate Change 23, 113-143.
Nonaka, I., 1991. The knowledge-creating company. Harvard business review 69, 96-104.
Nonaka, I., 1994. A dynamic theory of organizational knowledge creation. Organization Science 5, 14-37.
Owen-Smith, J., Powell, W.W., 2003. The expanding role of university patenting in the life sciences: assessing the importance of experience and connectivity. Research Policy 32, 1695-1711.
Polanyi, M., 1975. Personal knowledge, in: Polanyi, M., Prosch, H. (Eds.), Meaning. University of Chicago Press, Chicago, IL, pp. 22 - 45.
Régibeau, P., Rockett, K., 2010. Innovation Cycles and Learning at the Patent Office: Does the early patent get the delay? The Journal of Industrial Economics 58, 222-246.
Roach, M., Cohen, W.M., 2013. Lens or Prism? Patent Citations as a Measure of Knowledge Flows from Public Research. Management Science 59, 504-525.
Rosenberg, N., 1990. Why do firms do basic research (with their own money)? Research Policy 19, 165-174.
Sampat, B.N., Ziedonis, A.A., 2004. Patent Citations and the Economic Value of Patents: A Preliminary Assessment, in: Moed, H.F., Glänzel, W., Schmoch, U. (Eds.), Handbook of Quantitative Science and Technology Research. Kluwer Academic Publishers, Dordrecht.
Sapsalis, E., van Pottelsberghe de la Potterie, B., 2007. The institutional sources of knowledge and the value of academic patents. Economics of Innovation and New Technology 16, 139-157.
Schmoch, U., 2008. Concept of a Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Organisation (WIPO).
Sneed, K.A., Johnson, D.K., 2009. Selling ideas: the determinants of patent value in an auction environment. R&D Management 39, 87-94.
Sorenson, O., Fleming, L., 2004. Science and the diffusion of knowledge. Research Policy 33, 1615-1634.
Spencer, J.W., 2001. How relevant is university-based scientific research to private high-technology firms? A United States–Japan comparison. Academy of Management Journal 44, 432-440.
Stremersch, S., Verniers, I., Verhoef, P.C., 2007. The quest for citations: Drivers of article impact. Journal of Marketing 71, 171-193.
Subramanian, A.M., Lim, K., Soh, P.-H., 2013. When birds of a feather don’t flock together: different scientists and the roles they play in biotech R&D alliances. Research Policy 42, 595-612.
Teece, D.J., 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. Research Policy 15, 285-305.
Thompson, P., 2006. Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations. Review of Economics and Statistics 88, 383-388.
Thursby, J.G., Jensen, R., Thursby, M.C., 2001. Objectives, characteristics and outcomes of university licensing: A survey of major US universities. The Journal of Technology Transfer 26, 59-72.

Thursby, J.G., Kemp, S., 2002. Growth and productive efficiency of university intellectual property licensing. Research Policy 31, 109-124.
Tijssen, R.J.W., 2002. Science dependence of technologies: evidence from inventions and their inventors. Research Policy 31, 509-526.
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, 19-50.
Wooldridge, J.M., 2009. Introductory econometrics: A modern approach. South-Western Pub.
Ziedonis, R.H., 2004. Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. Management Science 50, 804-820.
Zollo, M., Winter, S.G., 2002. Deliberate learning and the evolution of dynamic capabilities. Organization Science 13, 339-351.
Zucker, L.G., Darby, M.R., Brewer, M.B., 1998. Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises. American Economic Review 88, 290-306.

Table 1. Descriptive statistics

	Variables	Number of Observations	Mean	Std Dev	Median	Mode	Minimum	Maximum
Dependent variable	Year group adjusted number of forward patent citations (STDCITECOUNT)	8908	1	1.98	0.36	0	0	29.84
Continuous variables	Average impact factor of journal articles cited in firm patents (IMPACT)	8908	9.42	5.49	8.47	5.38	0.23	50.81
	Average number of citations of journal articles cited in firm patents (ARTCITE)	8908	174.61	459.40	74.47	0	0	24929
	Number of cited journal articles (JARTs)	8908	25.21	35.01	11	0	0	189
	Number of backward patent references (PATREFS)	8908	37.38	61.93	16	0	0	749
	Average number of citations to referenced patents (REFQUAL)	8908	97.40	144.02	53.77	0	0	3447
	Number of IPC categories (SCOPE)	8908	3.31	2.02	3	2	1	17
	Originality index (ORIGINAL)	8908	0.61	0.25	0.64	1	0	1
	Number of inventors (INVENT)	8908	3.51	2.47	3	2	1	27
	Age of firm in years at patent application year (FIRMAGE)	8908	12.91	6.06	12	10	1	51
	Age of patent in months as of December 2012 (PATAGE)	8908	99.96	51.77	100	79	0	318
	Firm Size (FIRMSIZE)	8856	9.17	2.90	10	12	1	12
	Firm Average Article Citations (AVGCITE)	8908	45.91	46.96	44.97	0	0	279.53
	Firm articles published prior to patent applications (FIRMPUBS)	8908	122.80	228.53	19	0	0	1366
	Firm Research Level (RESLEV)	8908	2.27	1.58	3.24	0	0	4
Binary variable	Existence of university collaborator[a] (COLLAB)	303						
Other	Unadjusted number of forward patent citations	8908	15.47	40.68	4	0	0	904
	Firm size[b]	434	5.11	3.24	4	2	1	12

[a]For this variable, the number of observations lists the number of patents with at least one university as a co-assignee.

[b]This measure is the size of the individual, unique firms represented in our database. It is coded as follows: 1=1-4 employees, 2=5-9 employees, 3=10-14 employees, 4=15-19 employees, 5=20-24 employees, 6=25=49 employees, 7=50-74 employees, 8=75-99 employees, 9=100-149 employees, 10=150-249 employees, 11=250=499 employees, 12=500 or more employees.

Table 2. Correlation matrix

	STDCITECOUNT	IMPACT	ARTCITE	JARTS	PATREFS	REFQUAL	SCOPE	ORIGINAL	INVENT	PATAGE	COLLAB	FIRMAGE	FIRMSIZE	AVGCITE	FIRMPUBS
IMPACT	0.01	1													
ARTCITE	0.01	0.16	1												
JARTS	0.12	0.06	0.09	1											
PATREFS	0.29	-0.05	0.00	0.44	1										
REFQUAL	0.13	0.12	0.08	0.09	0.03	1									
SCOPE	0.06	0.09	0.03	0.15	0.12	0.07	1								
ORIGINAL	0.05	0.04	0.00	0.02	0.06	-0.02	0.11	1							
INVENT	0.07	-0.03	0.00	0.09	0.12	-0.05	0.03	0.01	1						
PATAGE	0.04	0.03	0.00	-0.21	-0.26	0.12	0.03	-0.01	-0.18	1					
COLLAB	-0.03	0.05	0.03	0.06	-0.03	0.00	-0.01	-0.02	0.08	-0.01	1				
FIRMAGE	-0.13	-0.05	-0.03	0.05	0.01	-0.06	-0.02	-0.04	0.10	-0.47	-0.04	1			
FIRMSIZE	0.03	0.08	0.03	-0.06	-0.03	0.12	0.08	0.01	0.03	0.13	-0.07	0.22	1		
AVGCITE	0.00	0.19	0.09	0.05	0.01	0.15	0.12	0.06	0.02	-0.01	0.01	0.14	0.42	1	
FIRMPUBS	-0.09	0.05	0.02	-0.01	-0.06	-0.03	0.00	-0.02	0.01	-0.22	0.00	0.28	0.42	0.29	1
RESLEV	-0.02	0.09	0.02	0.06	0.00	0.11	0.06	0.01	0.02	0.02	-0.05	0.22	0.47	0.67	0.37

Table 3. Tobit estimates. The dependent variable is year group adjusted number of forward patent citations.

Variables / Model	1		2		3		4		5		6		7		8	
	Estimate		Estimate		Estimate		Estimate		Estimate		Estimate		Estimate		Estimate	
Intercept	0.077		-1.390	***	0.432	**	-0.861	***	-0.734	***	-0.853	***	-0.883	***	-0.759	***
Average impact factor of journal articles cited in firm patents (IMPACT)	0.015	***	0.012	***	0.008	**	0.010	**	-0.003		0.010	**	0.010	**	-0.003	
Average number of citations of journal articles cited in firm patents (ARTCITE)	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
Number of cited journal articles (JARTs)	0.004	***	0.003	***	0.006	***	0.003	***	0.003	***	0.003	***	0.004	***	0.004	***
Number of backward patent references (PATREFS)			0.004	***			0.004	***	0.004	***	0.004	***	0.004	***	0.004	***
Average number of citations to referenced patents (REFQUAL)			0.001	***			0.001	***	0.001	***	0.001	***	0.001	***	0.001	***
Number of IPC categories (SCOPE)			0.002				-0.001		-0.001		-0.002		0.000		-0.001	
Originality index (ORIGINAL)			0.068				0.033		0.027		0.032		0.031		0.026	
Number of inventors (INVENT)			0.080	***			0.079	***	0.079	***	0.079	***	0.079	***	0.080	***
Age of patent in months as of December 2012 (PATAGE)			0.011	***			0.008	***	0.008	***	0.008	***	0.008	***	0.008	***
University collaborator (COLLAB)					-0.346	***	-0.386	***	-0.370	***	-0.370	***	-0.398	***	-0.374	***
Age of firm in years at patent application year (FIRMAGE)					-0.093	***	-0.048	***	-0.048	***	-0.047	***	-0.047	***	-0.048	***
Firm Size (FIRMSIZE)					0.094	***	0.047	***	0.045	***	0.044	***	0.046	***	0.045	***
Firm Average Article Citations (AVGCITE)					0.005	***	0.005	***	0.005	***	0.006	***	0.005	***	0.005	***
Firm articles published prior to patent applications (FIRMPUBS)					0.000	*	0.000		0.000		0.000		0.000		0.000	
Firm Research Level (RESLEV)					0.014		-0.018		-0.062		-0.020		-0.011		-0.056	
IMPACT * RESLEV									0.005	**					0.005	**
ARTCITE * RESLEV											0.000				0.001	
JARTS * RESLEV													0.000		0.000	
Firm binary variables included	YES		YES		YES		YES		YES		YES		YES		YES	
Tech class binary variables included	YES		YES		YES		YES		YES		YES		YES		YES	
Wald Test of Firm Binary Variables	2871.00	***	2487.00	***	2748.60	***	2416.80	***	2413.00	***	2411.80	***	2347.00	***	2333.90	***
Wald Test of Tech Class Binary Variables	230.93	***	152.04	***	218.98	***	153.75	***	152.80	***	155.64	***	153.65	***	153.39	***
Wald Test of Full Model	4110.20	***	4785.00	***	4520.70	***	4839.50	***	4851.20	***	4837.40	***	4835.60	***	4835.20	***
AIC	31252		30652		30714		30448		30446		30450		30450		30450	
-2 Log Likelihood	31062		30450		30512		30234		30230		30234		30234		30230	
Multicollinearity Index	8.31		15.95		24.72		35.99		38.89		34.08		36.76		40.57	
Number of Observations	8908		8908		8856		8856		8856		8856		8856		8856	

*** .01 significance, ** .05 significance, * .10 significance

Table 4. Robustness checks. Tobit estimates. The dependent variable is year group adjusted number of forward patent citations.

Variables / Model	1		2		3		4	
	Estimate		Estimate		Estimate		Estimate	
Intercept	-0.761	***	-0.975	***	-0.824	***	-0.814	***
Average impact factor of journal articles cited in firm patents (IMPACT)	-0.003		-0.003		-0.010			
Average number of citations of journal articles cited in firm patents (ARTCITE)			0.000		0.000			
ARTCITE - Year Group Adjusted	0.004							
Number of cited journal articles (JARTs)	0.004	***	0.004	***	0.004	**	0.004	***
Number of backward patent references (PATREFS)	0.004	***	0.004	***	0.005	***	0.004	***
Average number of citations to referenced patents (REFQUAL)	0.001	***	0.001		0.001		0.001	***
Number of IPC categories (SCOPE)	0.000		-0.004		0.031	***	0.004	
Originality index (ORIGINAL)	0.028		0.045		0.003		0.038	
Number of inventors (INVENT)	0.079	***	0.080	***	0.080	***	0.079	***
Age of patent in months as of December 2012 (PATAGE)	0.008	***	0.008	***	0.008	***	0.008	***
University collaborator (COLLAB)	-0.385	***	-0.381	***	-0.377	***	-0.373	***
Age of firm in years at patent application year (FIRMAGE)	-0.048	***	-0.048	***	-0.050	***	-0.048	***
Firm Size (FIRMSIZE)	0.046	***	0.042	***	0.029	*	0.051	***
Firm Average Article Citations (AVGCITE)	0.005	***	0.005	***	0.005	***	0.006	***
Firm articles published prior to patent applications (FIRMPUBS)	0.000		0.000		0.000		0.000	
Firm Research Level (RESLEV)	-0.058		-0.045		-0.062		-0.011	
IMPACT * RESLEV	0.005	**	0.005	**	0.007	**		
ARTCITE * RESLEV	0.001		0.001		0.003			
JARTS * RESLEV	0.000		0.000		0.000		0.000	
Firm binary variables included	YES		YES		YES		YES	
Tech class binary variables included	YES		YES		NO		YES	
Application year binary variables included	NO		YES		YES		NO	
Wald Test of Firm Binary Variables	2335.60	***	2290.30	***	2560.20	***	2338.40	***
Wald Test of Tech Class Binary Variables	149.59	***	142.46	***			150.95	***
Wald Test of Application Year Binary Variables			36.29	**	42.51	***		
Wald Test of Full Model	4840.20	***	3792.90	***	3766.70	***	4821.40	***
AIC	30450		30457		30559		30453	
-2 Log Likelihood	30230		30194		30336		30240	
Multicollinearity Index	37.78		40.57		37.64		32.31	
Number of Observations	8856		8856		8856		8856	

*** .01 significance, ** .05 significance, * .10 significance

Appendix Table 1

	Variable	Number of Observations	Mean	Std Dev	Median	Mode	Minimum	Maximum
Dependent variable	Year group adjusted number of forward patent citations (STDCITECOUNT)	15505	1.00	2.04	0.00	0	0	77
Continuous variables	Number of backward patent references (PATREFS)	15505	32.79	57.71	13.00	0	0	749
	Number of IPC categories (SCOPE)	15505	3.03	1.88	3.00	2	1	17
	Originality index (ORIGINAL)	15505	0.59	0.27	0.63	1	0	1
	Firm - University Collaboration Rate	15505	0.03	0.09	0.00	0	0	1
	Number of inventors (INVENT)	15505	3.34	2.39	3.00	2	1	33
	Age of firm in years at patent application year (FIRMAGE)	15505	12.93	6.97	8.00	6	0	96
	Age of patent in months as of December 2012 (PATAGE)	15505	107.49	58.85	107.00	122	0	473
Other	Firm size, unweighted[a]	910	4.78	3.11	4	2	1	12
	Age of firm in years at first patent application	910	7.49	6.17	6	5	1	93
	Unadjusted number of forward patent citations	15505	16.49	38.12	4	0	0	904

[a]This measure is the size of the individual, unique firms represented in our database. It is coded as follows: 1=1-4 employees, 2=5-9 employees, 3=10-14 employees, 4=15-19 employees, 5=20-24 employees, 6=25=49 employees, 7=50-74 employees, 8=75-99 employees, 9=100-149 employees, 10=150-249 employees, 11=250=499 employees, 12=500 or more

Appendix Table 2. Tobit estimates, outliers ommitted. The dependent variable is year group adjusted number of forward patent citations.

Variables / Model	1		2		3		4		5		6		7		8	
	Estimate		Estimate		Estimate		Estimate		Estimate		Estimate		Estimate		Estimate	
Intercept	0.050		-1.403	***	0.326	***	-0.903	***	-0.719	***	-0.896	***	-0.929	***	-0.748	***
Average impact factor of journal articles cited in firm patents (IMPACT)	0.022	***	0.019	***	0.015	***	0.016	***	-0.003		0.016	***	0.016	***	-0.004	
Average number of citations of journal articles cited in firm patents (ARTCITE)	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
Number of cited journal articles (JARTs)	0.004	***	0.004	***	0.006	***	0.004	***	0.004	***	0.004	***	0.004	***	0.005	***
Number of backward patent references (PATREFS)			0.004	***			0.004	***	0.004	***	0.004	***	0.004	***	0.004	***
Average number of citations to referenced patents (REFQUAL)			0.001	***			0.001	***	0.001	***	0.001	***	0.001	***	0.001	***
Number of IPC categories (SCOPE)			-0.011				-0.013		-0.011		-0.013		-0.011		-0.009	
Originality index (ORIGINAL)			0.102				0.057		0.050		0.057		0.064		0.055	
Number of inventors (INVENT)			0.074	***			0.072	***	0.072	***	0.072	***	0.072	***	0.072	***
Age of patent in months as of December 2012 (PATAGE)			0.011	***			0.008	***	0.008	***	0.008	***	0.008	***	0.008	***
University collaborator (COLLAB)					-0.353	***	-0.373	***	-0.361	***	-0.368	***	-0.377	***	-0.373	***
Age of firm in years at patent application year (FIRMAGE)					-0.087	***	-0.047	***	-0.048	***	-0.047	***	-0.047	***	-0.048	***
Firm Size (FIRMSIZE)					0.099	***	0.048	***	0.048	***	0.047	***	0.049	***	0.050	***
Firm Average Article Citations (AVGCITE)					0.005	***	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***
Firm articles published prior to patent applications (FIRMPUBS)					-0.001	**	0.000		0.000		0.000		0.000		0.000	
Firm Research Level (RESLEV)					0.021		-0.015		-0.089	**	-0.017		-0.007		-0.080	*
IMPACT * RESLEV									0.009	***					0.009	***
ARTCITE * RESLEV											0.000				0.000	
JARTS * RESLEV													0.000		0.000	
Firm binary variables included	YES		YES		YES		YES		YES		YES		YES		YES	
Tech class binary variables included	YES		YES		YES		YES		YES		YES		YES		YES	
Wald Test of Firm Binary Variables	2623.20	***	2271.80	***	2457.20	***	2180.20	***	2172.70	***	2179.30	***	2114.30	***	2101.20	***
Wald Test of Tech Class Binary Variables	234.51	***	160.14	***	222.28	***	161.72	***	158.44	***	161.61	***	161.81	***	158.15	***
Wald Test of Full Model	3802.20	***	4346.70	***	4156.10	***	4454.40	***	4471.60	***	4435.90	***	4471.90	***	4459.60	***
AIC	26232		25759		25771		25558		25551		25560		25560		25554	
-2 Log Likelihood	26048		25562		25574		25350		25340		25350		25350		25340	
Multicollinearity Index	8.05		15.88		22.78		30.02		32.56		30.28		30.74		33.48	
Number of Observations	7349		7349		7297		7297		7297		7297		7297		7297	

*** .01 significance, ** .05 significance, * .10 significance