Paper to be presented at the

DRUID Society Conference 2014, CBS, Copenhagen, June 16-18

The Division of Innovative Labor between Universities and Firms: Evidence from ?Knowledge Twins?

Michaël Bikard
London Business School
Strategy & Entrepreneurship
mbikard@london.edu

Abstract

Do firms and universities play distinct roles in the process of science-based invention and if so, how do they differ? This paper examines the relative impact of both types of organizations on knowledge use empirically. We disentangle the marginal impact of the organizational environment from the impact of selection, in which firms gravitate around scientific knowledge that has a higher technological potential. Exploiting simultaneous discoveries as ?knowledge twins,? we find that firms and universities are conducive to a different uses of scientific knowledge. While firms efficiently translate knowledge into patented inventions, universities tend to use it to produce yet more scientific knowledge in the form of scientific publications. These differences are also apparent within discovery teams when observing collaborations between firms and universities.

Jelcodes:O32,-

The Division of Innovative Labor between Universities and Firms: Evidence from "Knowledge Twins"

ABSTRACT

Do firms and universities play distinct roles in the process of science-based invention and if so, how do they differ? This paper examines the relative impact of both types of organizations on knowledge use empirically. We disentangle the marginal impact of the organizational environment from the impact of selection, in which firms gravitate around scientific knowledge that has a higher technological potential. Exploiting simultaneous discoveries as "knowledge twins," we find that firms and universities are conducive to a different uses of scientific knowledge. While firms efficiently translate knowledge into patented inventions, universities tend to use it to produce yet more scientific knowledge in the form of scientific publications. These differences are also apparent within discovery teams when observing collaborations between firms and universities.

Keywords:

Corporate R&D; Academic science; Inventions; Simultaneous discoveries; Patents

1. INTRODUCTION

Universities and firms both play important roles in the process of science-based invention. Entire industries as with semi-conductors or biotechnologies can track their origins to the benches of university labs. The division of innovative labor has been often described. Universities tend to focus on fundamental research whereas firms undertake more applied projects (Rosenberg and Nelson 1994; Mansfield 1998; Cohen, Nelson, and Walsh 2002; Sauermann and Stephan 2013). As the link between invention and public science intensifies (Narin, Hamilton, and Olivastro 1997; Branstetter 2005), the respective roles of firms and universities are however becoming less clear. On the one hand, firms are commonly involved in fundamental scientific research (Rosenberg 1990). On the other, a growing number of universities and public organizations undertake "translational activities" (Harris 2011). Taken together, these trends beg the question of the circumstances under which scientific research ought to be endeavored in a firm or in a university. Couldn't science-based invention occur entirely within firms (or within universities)? How does each type of organization shape scientific knowledge and its translation into new technologies?

Firms and universities certainly present different environments for scientific research. For example, firms are likely to under-invest in basic science because they tend to be unable to appropriate the economic value stemming from this type of work (Nelson 1959; Arrow 1962). They are more focused than universities, but they also incur higher costs, and will therefore tend to work on more applied projects (Aghion, Dewatripont, and Stein 2008). In addition, universities and firms have different missions, and this difference might impact the motivation of academic and corporate scientists (Lacetera 2009a). Although these studies have provided fascinating insights about the selection of firms and universities into different projects, little is known about the impact that both environments have on scientific knowledge. Yet the type of environment in which the research is done might not be inconsequential for the development of that knowledge. Using an in-depth case study, Stern (1995) has found that particularities of the academic environment had been obstacles to the rapid development of synthetic insulin at Harvard and UCSF. Absent a systematic examination of the impact that firms and universities have on scientific knowledge, the tradeoffs associated with the use of each type of organization remain unclear.

To answer this question, we consider that scientific knowledge might be used for the production of new technologies (Fleming and Sorenson 2004; Aghion, Dewatripont, and Stein 2008) but that it might also be used for the production of additional scientific knowledge (Furman and Stern 2011). In turn, this "dual usefulness" of scientific knowledge (Nelson 1962; Stokes 1997) provides a key to understanding the division of labor between universities and firms. Environments that are conducive to the translation of scientific knowledge might not be adapted to its use for further scientific exploration. In other words, a tradeoff might exist between organizing for using scientific knowledge as "shoulders" for additional research and organizing for using scientific knowledge as "map" fostering technology development. We propose that while universities are particularly conducive to using scientific knowledge as a springboard for further scientific explorations, firms tend to foster its translation into new technologies. Because both uses are ultimately beneficial to science-based invention, both firms and universities play fundamental but distinct roles in the division of innovative labor.

The challenge in exploring the uses of scientific knowledge empirically is considerable. The potential uses of a given piece of knowledge are always unobserved. Hence, measured rates of follow-on uses (e.g., follow-on publications or inventions) might result from the environment in which a given discovery was made, but it might also be a consequence of the nature and promise of that discovery. For instance, new knowledge produced in universities is likely to be more fundamental on average than scientific discoveries made by firms. In order to examine the relative impact of the academic and industry environments on science-based invention, it is crucial to account for the technological potential of the scientific discoveries. This paper reports a novel empirical strategy to tackle this challenge.

In the winter of 1999, two teams of scientists simultaneously discovered VR1 (vanilloid receptor-1), the receptor for the pain caused by excessive heat or capsaicin, the pungent component of chili peppers. The first team, led by Dr. John B Davis, sent its results to Nature on December 20, 1999 and the paper was published on May 11, 2000. The second team, led by Prof. David Julius, sent its results to Science on January 18, 2000 and the paper was published on April 14, 2000. The new knowledge had important implications for the development of pain therapeutics. Yet, both discoveries were made in very different organizations. Julius is an academic based at UC San Francisco. In contrast, Davis is an industrial scientist working at SmithKline Beecham. Simultaneous discoveries are a fascinating and relatively frequent phenomenon (Merton 1961). When the discoverers submit their findings for publication at almost the same time, two or more papers disclosing the same discovery can be accepted, thus leading to the publication of "paper-twins." Paper-twins are scientific articles that disclose the same underlying piece of knowledge. They are thus more than closely related or complementary discoveries. Rather, by embodying the same piece of knowledge that emerged in two distinct environments, paper-twins are a natural consequence of the duplication of effort in science, and a potentially rich setting to study the determinants of science-based invention.

The "experiment" afforded by the observation of discoveries occurring simultaneously in a university and in a firm will allow for a set of precise tests of the relative impact of both types of organizations on knowledge use. First, we examine the extent to which firm and university scientists translate the newly produced scientific knowledge into published scientific knowledge and patented inventions. Second, we examine the quality of follow-on publications and patents as measured by their citations. Third, we examine the marginal impact of both types of organizations on knowledge use in the context of university-firm collaborations. Because it enables the observation of the non-occurrence of patents and publications that could have occurred, paper-twins are a setting particularly suited to investigating the impact of different environments on knowledge use.

The analysis centers on 90 scientific papers disclosing thirty-nine simultaneous discoveries that involved at least one team from a university and another team from a firm. These 90 teams produced 171 patents and 320 publications based on the new knowledge, therefore allowing for a quantitative study of the impact of each organizational environment on knowledge use. Prior studies of the division of innovative labor have relied on formal modeling (Arrow 1962; Aghion, Dewatripont, and Stein 2008; Lacetera 2009a) or in-depth case studies (Stern 1995). Though our analysis centers on "only" 39 simultaneous discoveries, this is nevertheless the first quantitative investigation of the impact of the university and firm environments on the use of scientific knowledge. We complemented our econometric analysis by interviewing 21 scientists that were the corresponding author on at least one of the 90 publications. The results indicate that the dual usefulness of scientific knowledge lies at the core of the division of innovative labor between universities and firms. Keeping the discovery constant, scientists from industry produce over three times more patents than their colleagues in academia; but academic scientists produce over three times more publications than their industry counterparts. Our interviews further reveal that this robust difference in the organization's marginal impact stems from a number of factors including different missions and cost structures but not limited to them.

2. THE IMPACT OF INDUSTRY AND ACADEMIA ON NEW SCIENTIFIC KNOWLEDGE

2.1. A Division of Innovative Labor

The existence of a division of innovative labor between firms and universities has been well documented. In its simplest form, universities are known to focus on more fundamental questions and firms tend to use scientific knowledge to produce new technologies. These differences have been established empirically using both case studies and large scale surveys (Rosenberg and Nelson 1994; Mansfield 1998; Cohen, Nelson, and Walsh 2002; Sauermann and Stephan 2013). While the existence of a division of innovative labor is consensual, the mechanism underlying this division of innovative labor is not.

Universities and firms present very different organizational environments. In a nutshell, firms tend to focus on the private appropriation of the results of their research. In contrast, scientists at universities face a larger number of goals. Although university scientists often value economic impact, they need to balance this objective with their traditional pursuit of knowledge for knowledge's sake as well as with the educational mission of academic institutions. The widely studied difference in institutional logic (Merton 1973; Dasgupta and David 1994; Gittelman and Kogut 2003; Murray 2010) has important consequences concerning the daily work of scientists in both environments. Academic scientists enjoy greater freedom and are primarily rewarded for their contribution to the knowledge commons as apparent through their publication record. In contrast, corporate scientists tend to publish less and patent more—especially when they conduct work of more applied nature (Moon 2011; Sauermann and Stephan 2013). Firms also tend to offer lower levels of freedom but higher salaries (Stern 2004). These differences, in turn, attract different types of scientists: PhD students that select into academia have a stronger taste for the nonmonetary benefits of science (Roach and Sauermann 2010; Agarwal and Ohyama 2013). While these differences are likely to play a key role in the division of innovative labor, the nature of this role remains contested.

To better understand this division of innovative labor, a number of studies have focused on specific differences between firms and universities and have theorized about their implications using formal models. Historically, these models have focused on firms' difficulty to capture the economic value from their investment in research. Because competitive markets under-incentivize private investment in basic science, public support is justified (Nelson 1959; Arrow 1962). While they make a convincing case for subsidizing basic research, these studies provide little insight about whether these public subsidies should be given to universities or to firms. More recently, a few studies have departed from the traditional focus on appropriability and have instead emphasized firms' superior focus. Aghion, Dewatripont, and Stein (2008) explore the tradeoff between private sector focus and academic freedom. Because scientists value academic freedom, firms need to pay them a wage premium so that they give it up. The combination of higher wages but greater focus in firms makes it socially optimal for them to concentrate on later-stage research while academic scientists can conduct basic research at a lower cost. Lacetera (2009a; 2009b) also consider firms' focus. However, he does not contrast it with academic freedom but with academic scientists' multiple goals. In two models, he considers the impact of this difference in goals on the type of research and development activities that are pursued in firms and in universities. Lacetera (2009a) examines firms' decision to conduct an R&D project in-house or to outsource it to a university. By keeping an R&D project in-house, a firm keeps greater discretion over the project, and maintains its ability to terminate it at any time. This control comes at a cost, however, because scientists are aware that their project can be interrupted and therefore lose motivation. The model predicts that firm will outsource projects that have a broader applicability, but that they will keep longer projects in-house. Lacetera (2009b) focuses on the decision of scientists in industry and in academia to pursue commercial activities. He finds that the difference in goals of the scientists in both environments might lead academic scientists to more reluctantly move to commercially relevant activities, but that it might also drive them to commercialize faster, less finished projects than their counterparts in industry. Taken together, these studies have provided fascinating insights about the conditions under which universities and firms might respectively be used. However, they do have not considered that the two types of organizations might have a different impact on knowledge.

Yet, case studies have found that firms and universities might have a different impact on scientific knowledge. Stern (1995) explores how academic scientists' variety of goals affects their ability to produce science-based inventions efficiently using a rich description of the case of synthetic insulin in the late 1970s. Three teams played a key role in the expression of human insulin in E. coli bacteria. Two were primarily academic efforts: Walter Gilbert's group at Harvard University and William Rutter and Howard Goodman's groups at UCSF worked on this topic in parallel. The third effort was led by Herbert Boyer and Bob Swanson and was organized as a start-up called Genentech[1]. Genentech as an organization was focused on the development of synthetic insulin. In contrast, the UCSF and Harvard teams were trying to combine the pursuit

[1] All three teams involved collaborations between academia and industry. The organization of the team leader however differed, and this difference seems to have played a key role in the process of science-based invention in the case of synthetic insulin.

of the new technology with the development of new fundamental insights while training the labs' students. They chose to express insulin in bacteria using the DNA cloning (cDNA) method while Genentech used gene synthesis. The cDNA method promised richer fundamental insights but was much more difficult to implement than gene synthesis. In addition, Genentech made hiring decisions that were designed to increase the efficiency of the process of science-based invention whereas the academic teams faced tradeoffs between research speed and students' training. On August 24, 1978, the Genentech team successfully expressed human insulin in bacteria. At that time, the Harvard and UCSF teams had just started to seriously organize their commercial activity (Stern 1995). Similar patterns have been found in other case studies. For instance, academic orientation might also be reason why the transistor was invented at Bell Labs rather than at Purdue University (Bray 1997; Fini and Lacetera 2010).

Universities and firms, then, might have a different impact on the development of scientific knowledge. This impact, however, has received little attention beyond Stern's in-depth exploration. Yet, absent a clear understanding of this impact, the tradeoffs associated with conducting research in either environment remain little understood.

2.2. Two Different Uses of Scientific Knowledge

We propose that firms and universities are two types of organizations that are conducive to a different use of scientific knowledge. Science can potentially be used in two ways, and firms and universities are each best adapted to one of these two uses.

Scientific knowledge can be developed with two possible goals in mind. Nelson (1962) for example noted that the invention of the transistor was motivated by the hope of producing both scientific and practical advances. Stokes (1997) distinguished between research that is inspired by consideration of use and research that is inspired by a quest for fundamental understanding. This potential “dual usefulness” of scientific knowledge stems from the nature of scientific knowledge itself. By uncovering new natural phenomena and regularities, scientific discoveries can potentially open two doors. On the one hand, it can open the door toward the development of a set of instructions to use these phenomena or regularities for a purpose (Mokyr 2002). In particular, new scientific knowledge might provide guidance in the process of invention, thereby vastly decreasing its cost (Nelson 1982; Fleming and Sorenson 2004). On the other hand, scientific discoveries can also open the door to new efforts toward a more detailed fundamental understanding of those phenomena and regularities. Indeed, science as an institution is cumulative, and each discovery can in principle provide new “shoulders” on which future research will stand (Merton 1973; Furman and Stern 2011).

This potential dual usefulness of scientific knowledge has deep implications for researchers and policy-makers because different types of environments foster different types of uses. For example, concerns have been raised that intellectual property rights, while fostering science-based invention, might at the same time deter cumulative scientific research. A number of researchers have highlighted the benefits of IP rights for science-based invention, emphasizing the creation of a market for ideas as well as its role in fostering investment in knowledge that has commercial potential (Kitch 1977; Hellmann 2007). On the other hand, recent research has found that IP rights might have a negative impact on follow-on scientific research. Using patent-paper pairs, Murray and Stern (2007) have found that the granting of a patent protecting the content of a recently published paper decreases follow-on citations to the paper by 10-20%. Conversely, Murray et al. (2011) have found that an NIH-induced reduction in access cost to patented engineered mice increases follow-on research and encourages the exploration of more diverse research paths. Interestingly the debate about the impact of IP rights on the two uses of scientific knowledge is not close. Williams (2013) recently found that patents on gene sequences have had a negative impact on both types of uses. She compares two types of gene sequences: those genes that were first sequenced by the Human Genome Project stayed in the public domain whereas those that were first sequenced by the firm Celera were temporarily protected by intellectual property. Strikingly, Williams finds that Celera's IP reduced subsequence research and product development on the order of 20-30%. The debate surrounding the impact of IP rights highlights the importance of considering the dual usefulness of scientific knowledge. Environments fostering invention are not necessarily beneficial to cumulative science, and vice versa.

We propose that firms are a type of organization that is best adapted to explore the technological potential of scientific knowledge. Firms' superior focus on economic value allows them to make more efficient decisions, both regarding hiring and research direction. More importantly perhaps, unlike their academic counterparts, firm scientists suffer no penalty from working on research projects that lack any scientific or educational value. In clear contrast, academic scientists suffer no penalty from working on projects that lack any economic value. Academic scientists are strongly incentivized to contribute to the scientific commons in order to further their own career. Certainly, at times, firms do publish and universities do patent (e.g., Rosenberg 1990; Henderson, Jaffe, and Trajtenberg 1998). However, on average, we propose that the marginal impact of firms and universities on scientific knowledge ought to be different. In short, firms tend to use a piece of knowledge to produce more inventions than universities whereas universities tend to use a piece of knowledge to produce more scientific publications than firms. We make the following predictions:

H1: Firm scientists are more likely to turn a specific scientific discovery into a larger number of patents than university scientists.

H2: University scientists are more likely to turn a specific scientific discovery into a larger number of scientific publications than firm scientists.

3. EMPIRICAL APPROACH

3.1. The Challenge

The empirical challenge in examining the treatment effect of different types of organizations on knowledge use is considerable. For example, when observing the emergence of science-based inventions, how can we gauge whether these stem from the intrinsic potential of the scientific knowledge itself or from the characteristics of the organization of discovery? Universities are widely believed to conduct much more basic research than firms. As a consequence, the relevance of university research for invention tends to be more indirect. The fundamental empirical challenge is therefore an identification problem. The risk is to conflate the marginal impact of the environment of discovery with the selection effect of knowledge into this environment. A simple comparison between different types of environments (e.g., university vs. industry) might therefore lead to biased results due to unobserved differences in the knowledge's potential. Ideally, the researcher would like to compare the (observed) knowledge use with the (unobserved) knowledge potential uses.

3.2. Paper Twins

This paper proposes a novel empirical approach exploiting the existence of simultaneous discoveries operationalized as paper twins. Paper twins are the dual instantiation of the same piece of new scientific knowledge in two distinct environments. The following example resulted from a discovery simultaneously made at UCSF and at SmithKline Beecham:

Caterina et al. (April 2000) "Impaired Nociception and Pain Sensation in Mice Lacking the Capsaicin Receptor." Science

"The capsaicin (vanilloid) receptor VR1 is a cation channel expressed by primary sensory neurons of the "pain" pathway. Heterologously expressed VR1 can be activated by vanilloid compounds, protons, or heat (>43°C), but whether this channel contributes to chemical or thermal sensitivity in vivo is not known. Here, we demonstrate that sensory neurons from mice lacking VR1 are severely deficient in their responses to each of these noxious stimuli. VR1−/− mice showed normal responses to noxious mechanical stimuli but exhibited no vanilloid-evoked pain behavior, were impaired in the detection of painful heat, and showed little thermal hypersensitivity in the setting of inflammation. Thus, VR1 is essential for selective modalities of pain sensation and for tissue injury–induced thermal hyperalgesia."

Davis et al. (May 2000) "Vanilloid receptor-1 is essential for inflammatory thermal hyperalgesia." Nature

"The vanilloid receptor-1 (VR1) is a ligand-gated, non-selective cation channel expressed predominantly by sensory neurons. VR1 responds to noxious stimuli including capsaicin, the pungent component of chilli peppers, heat and extracellular acidification, and it is able to integrate simultaneous exposure to these stimuli (…). Here we have disrupted the mouse VR1 gene using standard gene targeting techniques. (…) Although the VR1-null mice appeared normal in a wide range of behavioural tests, including responses to acute noxious thermal stimuli, their ability to develop carrageenan-induced thermal hyperalgesia was completely absent. We conclude that VR1 is required for inflammatory sensitization to noxious thermal stimuli but also that alternative mechanisms are sufficient for normal sensation of noxious heat."

These excerpts describe two sets of results obtained by examining the behavior of mice lacking a specific receptor (VR1). Both teams have found that mice in which the VR1 gene had been disrupted exhibit normal reactions to a variety of stimuli but become completely insensitive to one specific stimulus (carrageenan-induced thermal hyperalgesia). One of the team (Caterina et al.) conducted its research within academia and the other team (Davis et al.) in a firm. Both papers were submitted within a month (respectively, January 18th 2000 and December 20th 1999). In short, the (nearly) simultaneous discovery of the capsaicin receptor in two different environments led to the disclosure of the same new knowledge in two distinct papers.

We use simultaneous discoveries as an "experiment" from which it is possible to compare the relative impact of the academic and corporate environments on knowledge use. Specifically, our empirical strategy exploits three key aspects of the phenomenon associated with the production of paper-twins: (a) since they disclose the same discovery, the knowledge disclosed in each of the paper-twins has intrinsically the same potential for follow-on use; (b) since simultaneous discoveries emerge in different organizations, the knowledge from each discovery might not actually be used in the same way; (c) knowledge use in different organizations can be measured by observing publications and patents citing the twin-papers.

4. DATA AND METHODS

4.1. Sample definition

The data for this study is based on the first automatically and systematically collected dataset of simultaneous discoveries. The full dataset consists in 1,246 papers disclosing 578 discoveries published between 1970 and 2009. The core of the analysis presented in this paper is, however, based on a subset consisting of 90 scientific publications disclosing 39 simultaneous discoveries having involved at least one industry-based team and one team based in a public research organization. The algorithm used to build this dataset is based on the insight that two papers disclosing the same simultaneous discovery are systematically cited together in the follow-on scientific literature, not only in the same papers, but also in the same parenthesis, or adjacently (Cozzens 1989). Figure 1 summarizes the algorithm[2].

Insert Figure 1 about here

Our data is drawn from several sources. Data about each publication comes from ISI Web of Science, Scopus and Pubmed. Details about the corresponding authors and corresponding organization come from an analysis of the text of the publications. Follow-on patents and

[2] Tests of within-twin similarity omitted in AOM submission because of space constraints

publications data (through October 2013) were collected using Scopus and Pubmed respectively. Table 1 provides a list of variables and definitions.

Insert Table 1 about here

4.2. Dependent Variables[3]

Follow-on Inventions: We measure NPAT, the extent to which an organization uses a new piece of scientific knowledge in order to produce inventions, by counting the number of patents originating from one of the discovery organizations and that followed each simultaneous discovery. In order to examine if the difference in quantity of follow-on invention could conceal different patenting standards, we also considered a citation-weighted patent count.

Follow-on Publications: We measure NPUB, the extent to which an organization uses a piece of scientific knowledge in order to deepen a given line of research, by counting the number of publications originating from the discovery organizations and that followed each simultaneous discovery. Just like with patents, we also examine a citation-weighted publication count.

4.3. Independent Variable – Type of Organization of Discovery

Our main explanatory variable, ACADEMIA, is a dummy variable that equals to one if the organization is a university or public research organization and zero if it is a firm[4].

4.4. Control Variables

Knowledge*'s Potential*: Our main empirical strategy to account for the unobserved potential uses of scientific knowledge is to focus our analysis on simultaneous discoveries. In our analysis, PAPER TWIN is a fixed effect for each simultaneous discovery. In addition, we might

[3] Details about variable construction were omitted in AOM submission because of space constraints

[4] Since both types of organizations are of interest here we could just as well have considered an indicator variable called INDUSTRY=1-ACADEMIA

expect the relative impact of the university and firm environments to depend on the potential uses of the knowledge. Our proxy for a discovery's technological (scientific) potential is the discovery's yearly rate of citation in the patent (publication) literature by organizations that were not involved in the discovery.

Patent-Paper Pair: In order to account for the existence of patent pairs automatically and systematically, we developed a variant of the method used by Thompson and colleagues (Thompson, Mowery, and Ziedonis 2013). Our systematic method to find patent twins to papers is detailed in Figure 2[5]. Empirically, we created an indicator variable PATENT PAIR that takes the value one if the discovery organization has a patent pair and zero if it does not.

Insert Figure 2 about here

US-based Organization: We included an indicator variable that takes the value of one if the organization is based in the US (using the address figuring on the publication).

4.5. Empirical Analysis[6]

We study the use of scientific knowledge by organization i that take part to simultaneous discovery j. To do so, we examine the rate at which each of the discovery organizations generates patents and/or publications based on the new knowledge.

Main Test: To test our key hypotheses we focus on the organization of the corresponding authors on each twin paper[7]. Empirically, measuring follow-on use through patents and publications implies that we must account for its form as count data skewed to the right, calling for a count model such as a fixed-effect Poisson with quasi-maximum likelihood (i.e., "robust")

[5] Details about this algorithm were omitted in AOM submission because of space constraints
[6] Details about the empirical analysis were omitted in AOM submission because of space constraints
[7] In 7 papers, the corresponding author had both an academic and an industry affiliation. For these ambiguous cases, we have used the author's email address in order to assign an organization type. In addition, we also drop the simultaneous discoveries that they were involved with in our robustness analysis.

estimates. Our test for the relative impact of the academic and industry environments on knowledge use by the discovery organization i of twin j is:

$$USE_{ij} = f(\varepsilon_{ij};\ \alpha ACADEMIA_i + PAPER\ TWIN_j + \beta PATENT\ PAIR_{ij} + \delta X_i)$$

where USE_{ij} is either $NPAT_{ij}$ or $NPUB_{ij}$ the number of patents and publications produced by the discovery organization that build on the simultaneous discovery. As indicated above, we use both a simple count and a citation-weighted count of papers and patents. Our main explanatory variable, $ACADEMIA_i$, is an indicator that takes the value one if organization i is a university or a research institute and zero if it is a firm. $PAPER\ TWIN_j$ is our simultaneous discovery-level fixed effect and $PATENT\ PAIR_{ij}$ is an indicator variable that equals to one if organization i was awarded a patent pair to the discovery j. Finally, X_i is a vector of control variables including for instance a dummy variable if organization i is based in the US.

In order to gain a yet finer understanding of the relative impact of the university and firm environments on knowledge use, we further explored (a) whether there are systematic differences in the quality of the follow-on patents and publications as measured through citations and (b) whether similar patterns of follow-on use can be observed within-paper by focusing on academic-industry collaborations.

Quality of Follow-on Papers and Patents: Diverging rates of patenting and publishing could conceal differences in the quality of the patents and publications produced by firms and universities. We investigate the existence of such difference in our data by exploring the impact of the firm and university environments on the number of citations received by their follow-on papers and patents. For this analysis, each follow-on paper (patent) constitutes one observation. We use an OLS regression with robust standard errors clustered at the level of the paper-twin and estimate for each publication (patent) k of organization i following simultaneous discovery j:

$$\ln(1 + CITES_{ijk}) = f(\varepsilon_{ijk};\ \alpha ACADEMIA_i + PAPER\ TWIN_j + \delta X_{ik})$$

where $CITES_{ijk}$ is either PUB_CITES_{ijk} or PAT_CITES_{ijk} the number of citations received by follow-on publications (or patents). As before, $ACADEMIA_i$ is our main explanatory variable, $PAPER\ TWIN_j$ is our simultaneous discovery-level fixed effect, and X_{ik} is a vector or control variables including the publication year (filing year) of publication (patent) k and whether organization i is US-based.

University-Firm Collaborations: In order to ensure that our results are indeed revealing of the broader division of innovative labor, we explore follow-on use by universities and firms that collaborated on the same discovery paper. Empirically, we no longer study only the corresponding organization of each paper. Instead, we focus on each address figuring on the subset of our 1246 twin papers that were university-firm collaborations. We then test the relative impact of the university and firm environments on knowledge use by address m of paper l:

$$USE_{ml} = f(\varepsilon_{ml};\ \alpha ACADEMIA_m + PAPER_l + \beta PATENT\ PAIR_{ml} + \delta X_m)$$

where USE_{ml} is either $NPAT_{ml}$ or $NPUB_{ml}$ the number of patents and publications produced by the discovery organization figuring in address m that builds on discovery paper l. $ACADEMIA_m$ is our main explanatory variable, $PAPER_l$ is a publication-level fixed effect, $PATENT\ PAIR_m$ is a dummy variable equals to 1 if the organization of address m has been awarded a patent pair, and X_m is a vector of address-level control variables.

5. RESULTS

5.1. Sample

Sample Description: Our main analysis focuses on a sample of 90 publications disclosing 39 simultaneous discoveries that involved at least one firm paper and a university one. Table 2 describes our main variables. The oldest of the simultaneous discoveries in our data dates back from 1994 and the most recent occurred in 2008. 58 of the 90 teams (64%) were based in the US. On average, each discovery is used to produce 1.90 follow-on patents and 3.56 follow-on publications. 88% of the 171 follow-on patents and 69% of the 320 follow-on publications that we observe stem from a US-based organization. Our analysis of university-industry collaboration is based on 773 addresses from 131 collaborative papers. 58% of those addresses are US-based, and each address produced on average 0.17 follow-on patent and 1.43 follow-on publications.

Insert Table 2 about here

Out of our 90 publications, 49 have a corresponding author that was based in academia and 41 have a corresponding author that was based in a firm. Our analysis of within-twin similarity described in section 4.1 confirmed that those papers did indeed disclose the same discovery. Though within-twin citation difference ought to be small considering our algorithm, this difference could nevertheless be systematically correlated with the type of organization that produces the paper. Figure 3 shows that we do not find such a correlation in our data. A Wilcoxon-Mann-Whitney test shows that there is no statistically significant difference between the yearly citation rates of our university and firm papers ($z = 0.37$; $p = 0.71$).

Insert Figure 3 about here

Sample in Perspective: Our 39 simultaneous discoveries are likely to be scientifically more important than the average published discovery for three reasons. First, since we selected simultaneous discoveries based on co-citation patterns, our dataset includes only well-cited papers. Second, our interviews revealed that scientific journals often like to publish several papers disclosing the same discovery, especially when the discovery is important. Third, many of those simultaneous discoveries resulted from scientific races, and the latter are more likely to occur when a discovery appears particularly promising ex-ante. Overall, then, we should expect that the 39 discoveries in our data have a higher-than-usual scientific potential.

We should also expect that our 39 simultaneous discoveries have on average a high technological potential. Firms typically choose to work on knowledge that might have some technological implications. Since our simultaneous discoveries emerge in a firm at the same time as in a university, we are selecting on the type of project that firms undertake. Table 3 compares simultaneous discoveries that involve only university teams, those that involve only firms and our dataset of 39 simultaneous discoveries that involves both. Firm-firm twins produced 3.5 follow-on patents and 1.3 follow-on publications on average whereas and these numbers were 0.2 and 5.2 respectively for university-university twins. Of course, it is impossible to know if these large differences stem from fact that universities tend to work on more fundamental knowledge or from the fact that universities and firms tend to use knowledge differently. Interestingly, our dataset of university-firm twins generated an intermediate number of follow-on patents and publications with 1.9 follow-on patents and 3.6 follow-on publications.

Insert Table 3 and 4 about here

In order to put our 90 twin publications in perspective, we collected a sample of "regular" scientific publications. For each paper, we garnered information about the preceding article and the following article in the same issue of the same journal[8]. We compare the 90 twin papers of our sample with the 180 non-twins in Table 3. We find that our 90 papers received on average nearly 7 times as many patent citations and 3 times as many publication citations as the 180 papers that were adjacent to them. We further explore the shape of the distribution in Figure 4

[8] In practice, our script minimized the page differences and selected two articles of the same type in the same issue of the same journal

and find that the 90 papers in our sample have a much flatter distribution than the more traditional skewed distribution of the 180 adjacent articles. Clearly, the discoveries in our data display a much higher variance in terms of scientific and technological impact than typical academic papers and our dataset includes a disproportionate number of discoveries with high technological and/or high scientific impact.

Insert Figure 4 about here

In short, the 90 papers in our sample are not average scientific publications. Since we examine the use of scientific knowledge conditional on knowledge potential, our research setting ought to include discoveries with high technological and/or scientific potential. Even if they are common, scientific discoveries that have little or no potential are likely to stay unused whether they emerge in a firm or in a university. Because of the special characteristic of our dataset our results ought to be interpreted carefully. However, selecting on important discoveries has also advantages since these are presumably the discoveries that scientists, managers and policy-makers are most interested in.

5.2. Main Test

Table 5 and 6 present the main result of our analysis. They consider respectively follow-on invention and follow-on publication based on our dataset simultaneous discoveries. By allowing the observation of the same discovery in two distinct environments, simultaneous discoveries make it possible to identify systematic variance in knowledge use. Models (5-1) and (5-2) examine H1 and show that, conditional on a discovery's technological potential, it is developed into over three times more patents if it emerges in a firm as opposed to a university. Of course, quantity is not quality, and one could be concerned that firms might have lower patenting standard than universities. Models (5-3) and (5-4) shows that the effect is stronger if we consider citation-weighted patent count as opposed to a simple count. These results confirm our hypothesis that firms use scientific knowledge to produce more patented inventions.

The difference in science-based invention observed in Table 5 could stem from the fact that academic scientists tend to use scientific knowledge for a different purpose than translation into new technology. Table 6 examines the extent to which firms and universities use new knowledge as "shoulders" to produce yet more scientific publications. Confirming H2, models (6-1) and (6-2) show that university scientists use the new knowledge to produce over three times more new publications than firm scientists. Here too, we examine the possibility that firms and universities might have different publication standards by using a citation-weighted count of publications. Models (6-3) and (6-4) shows that our results remain essentially the same.

Figure 5 shows the different uses that universities and firms make of scientific knowledge graphically. The two plots show follow-on publishing (left-hand side) and follow-on patenting (right hand side) for firms (X-axis) and universities (Y-axis). Each dot corresponds to a paper twin. Since potential use is identical, if the type of discovery organization did not matter, we would expect the dots to stay close to the diagonal. Instead, we find that firms are much more likely to use the knowledge to produce more publications and firms are more likely to use it to produce more patents. Interestingly, the graph suggests that the effect is especially salient for knowledge that has high technological or high scientific potential. Indeed, the firms in our data rarely use the knowledge to produce more than five publications. Similarly, the universities in our data rarely use the knowledge to produce more than five patents.

Table 7 further estimates the robustness of our main result. In (7-1) and (7-4), we dropped simultaneous discoveries that involved more than two organizations and find that our result remain essentially the same. Models (7-2) and (7-5) exclude 7 twin papers for which we had one corresponding author that had both an academic and a corporate affiliation. As expected, the exclusion of these "hybrids" increases the magnitude of our main result and we find that firms produce more than four times more patents than universities and universities more than four times more publications than firms. If our results had been driven by a few outliers, we would have been able to see this in Figure 5. Models (7-3) and (7-6) nevertheless examines this possibility further by applying a 90% Winsorization to our dependent variables. The magnitude of our result decreases but the effect remains visible.

In order to more precisely understand our main result, table 8 examines how the effect of the organization of discovery varies with (a) the technological potential of the discovery, (b) the scientific potential of the discovery and (c) ownership of a patent pair. Indeed, one of the downside of using paper-twin fixed-effects is that it does not allow us to observe variation in our main effect across discoveries. In section 5.1, we argue that the effect of the organizational environment should only be visible for discoveries that are potentially useful. If a discovery has no potential for use, it is likely to remain unused whether it emerges in a firm or in a university. Models (8-1) and (8-4) examine how our main effect interacts with technological potential.

Model (8-1) focuses on patenting. As expected, in the absence of technological potential our coefficient for ACADEMIA becomes close to zero. No one patents. The impact, however, becomes significantly larger when technological potential increases, indicating that firms might seize technological opportunities more aggressively. Model (8-4) focuses on publishing and paints a slightly different picture. In the absence of technological potential, the difference in publishing rates between firms and universities is the highest. Interestingly, firms seem to increase their publication rate when the technological potential increases. These results are important because they show a possible connection between the type of use and the rate of use of new knowledge. While we mostly focus on the former, this result suggests that the rate of use of knowledge with low technological potential might be higher in universities whereas the rate of use of knowledge with high technological potential might be higher in firms. Because both types of knowledge exist, this result highlights the importance of the division of innovative labor between universities and firms.

Models (8-2) and (8-5) investigate the interaction with scientific potential. Before we examine the result, it is important to remember that while our dataset includes discoveries with no technological potential, it does not include discoveries with no scientific potential. All the discoveries in our data are not only published, they are also well cited, and our dataset therefore includes less variance with scientific potential than with technological potential. In fact, models (8-2) and (8-5) show non-statistically significant interaction effects between the organization of discovery and scientific potential. Finally, models (8-3) and (8-6) examine the interaction between the organization of discovery and having a paired patent. We find that that having a paired patent significantly increases our main effect with regard to patenting but not with regard to publishing.

5.3. Quality of Follow-on Papers and Patents

Table 9 examines quality of the follow-on papers and patents produced by universities and firms as measured by their respective citations. Such an analysis is important because firms and universities might have different “quality thresholds”. For instance, a number of studies have found that academic patents receive on average more citations than industry ones (Henderson, Jaffe, and Trajtenberg 1998; Mowery, Sampat, and Ziedonis 2002), therefore suggesting that they might be more selective when it comes to patenting. With first examine citations to publications. Models (9-1) to (9-3) show that there is no statistically significant difference between the number of citations received by follow-on publications, whether they stem from firms or academia. Models (9-4) to (9-6) examine the case of patent citations. Interestingly, we find that the average industry patent in our data receives significantly more citations than the average university patent. This result suggest that prior studies' result that firm patents receive on average fewer citations might be driven by differences in the nature of the inventions that are patented in both environments. In fact, model (9-6) shows that the negative impact of academia on patent citations disappears after we add the twin fixed-effects. This result has two implications. First, firms' patents might be more concentrated toward more promising lines of research than university patents (consistent with model 8-1). Second, within a given line of research, firm and university seem to generate patents of similar quality.

5.4. University-Firm Collaborations

Finally, since our argument is about the division of innovative labor between firms and university, we believe that we should be able to replicate our main finding using a separate dataset of university-firm collaborations. Of our 1,246 papers, 131 stemmed from such collaboration, therefore making it possible to study the division of innovative labor within papers. Specifically, we considered that firms and universities that made a joint discovery have admittedly the same knowledge at the same time—and therefore also the same potential for follow-on invention and publication. This test of our main hypothesis is inferior to the study of twins because the decision to get involved in follow-on research or invention is unlikely to be independent among collaborating organizations. However, if a division of innovative labor really exists, as we argued, we should be able to replicate our main finding with this different dataset.

Table 10 and 11 present the same analysis as our tables 5 and 6, although with a different dataset. The data is organized at the address level, with paper-level fixed effects. In line with our expectations, we are able to replicate our main finding. The collaborating university produces on average half the number of patents but far more publications (7 to 10 times more) than its corporate collaborator. The fact that our coefficient with this dataset are different than with the 39 simultaneous discoveries is interesting but not surprising. As we saw in table 8, the magnitude of the coefficients will depend on the technological potential of the discoveries that we are studying. Since our dataset is different from the one that we examined in table 5 and 6, there is no reason to think that we should have found coefficients of the same magnitude. In fact the smaller coefficient for patenting differences and higher coefficient for publishing differences suggests that the 131 collaborative papers have on average a lower technological potential than the 39 simultaneous discoveries that constitute the core of our analysis.

6. DISCUSSION AND CONCLUSIONS

Both firms and universities produce scientific discoveries; and both of them at times translate their discoveries into new technologies. Yet the decision to conduct research in one type of organization or the other is not inconsequential. Firms and universities, we argue, play distinct roles in the process of science-based invention. Instead of focusing on the selection of different projects in universities and firms, this paper investigates the marginal impact of each type of organization on new scientific knowledge. Studying this impact empirically is difficult because universities and firms tend to work on projects that have very different potential for scientific and technological developments. To address this identification challenge, we focus on simultaneous discoveries in science. By observing the same discoveries being made by different teams in firms and in universities around the same time, we are able to track the impact that each type of organization has on knowledge use.

We find that firms and universities use scientific knowledge for different purposes. The dual usefulness of scientific knowledge therefore lies at the core of the division of innovative labor between universities and firms. Universities tend to use new scientific knowledge in order to produce yet more new knowledge whereas firms tend to use it to produce new science-based inventions. Keeping the discovery constant, our data suggest that scientists from industry produce over three times more patents than their colleagues in academia. In contrast, academic scientists produce three times more publications than their industry counterparts. Taken together, these results suggest that conducting scientific research in one type of organization or the other is likely to bring different outcomes. By conducting research in an academic environment, managers and policy-makers are likely to gain fundamental understanding on a specific topic, and the latter is likely to be disclosed in the form of publications. In contrast, universities are not well adapted to translate scientific knowledge into new technologies, and firms are likely to be a more effective type of organization for this purpose.

This research is not without limitations: First, the simultaneous discoveries in our dataset occur in the same planet – i.e. although many of our twins occur on different continents, we cannot exclude the possibility of contagion whereby the decision to use knowledge in a certain way in one organization is influenced by the other teams' decisions. Second, the thirty-nine discoveries in our dataset that were made simultaneously in at least a firm and a university belong to a limited number of scientific specialties, especially immunology, oncology, and neuroscience, but also including materials science. For the majority of scientific disciplines including our own, no overlap exists between the scientific research that is conducted in industry and in academia, therefore raising questions about the generalizability of our conclusions. Third, we measure differences in knowledge use by examining publication and patent output. We are not able to examine whether these differences are driven by diverging disclosure choices or whether they are driven by different research directions. Nonetheless, our setting provides the advantage that we are able to distinguish empirically between the selection process whereby firms and universities tend to work on different topics and the marginal impact of each organizational environment on knowledge use. Moreover, our results appear robust to a number of different specifications and can be replicated by analyzing collaborative papers between academia and industry.

Our study is only a first step toward understanding the different roles that firms and universities play in the process of science-based invention. Other differences are likely to shape the decision to conduct research in one environment or the other, including the cost of labor in the two environments (Stern 2004), the varying levels of control that each environment offers over the research (Lacetera 2009a; Aghion, Dewatripont, and Stein 2008), as well as the differing social networks in which each type of organization is embedded (Murray 2002). These are all important elements that are likely to impact the division of labor between industry and academia. The importance of continuing the investigation of this division of labor should not be understated. As universities are increasingly involved in firm's innovation strategy, more attention might usefully be brought to the fact that in practice, firms and universities do not only produce different types of scientific knowledge, but that they also have a distinct impact on the knowledge that they produce.

REFERENCES

Agarwal, Rajshree, and Atsushi Ohyama. 2013. "Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists." Management Science 59 (4) (April 1): 950–970. doi:10.1287/mnsc.1120.1582.

Aghion, Philippe, Mathias Dewatripont, and Jeremy C. Stein. 2008. "Academic Freedom, Private-Sector Focus, and the Process of Innovation." The RAND Journal of Economics 39 (3) (October 1): 617–635.

Arrow, K. 1962. “Economic Welfare and the Allocation of Resources for Invention.” In R.R. Nelson (ed.), The Rate and Direction of Inventive Activity, 609–625. Princeton, NJ: Princeton University Press.

Bikard, Michaël. 2012. “Simultaneous Discoveries as a Research Tool: Method and Promise.” MIT Sloan Working Paper.

Bray, R. 1997. “A Case Study in Serendipity: Why Was the Transistor Invented at Bell Laboratories and Not at Purdue University?” The Electrochemical Society Interface, Spring: 24–31.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh. 2002. “Links and Impacts: The Influence of Public Research on Industrial R&D.” Management Science 48 (1): 1–23.

Cozzens, Susan E. 1989. Social Control and Multiple Discovery in Science: The Opiate Receptor Case. State Univ of New York Press.

Dasgupta, Partha, and Paul A. David. 1994. “Toward a New Economics of Science.” Research Policy 23 (5) (September): 487–521.

Fini, Riccardo, and Nicola Lacetera. 2010. “Different Yokes for Different Folks: Individual Preferences, Institutional Logics, and the Commercialization of Academic Research.” Advances in the Study of Entrepreneurship, Innovation & Economic Growth 21: 1–25.

Fleming, Lee, and Olav Sorenson. 2004. “Science as a Map in Technological Search.” Strategic Management Journal 25: 909–928.

Furman, Jeffrey, and Scott Stern. 2011. “Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research.” American Economic Review 101 (5): 1933–1963.

Gittelman, Michelle, and Bruce Kogut. 2003. “Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns.” Management Science 49 (4) (April 1): 366–382.

Harris, Gardiner. 2011. “Federal Research Center Will Help Develop Medicines.” The New York Times, January 22, sec. Health / Money & Policy.

Hausman, J., B.H. Hall, and Z. Griliches. 1984. “Econometric Models for Count Data with an Application to the Patents-R & D Relationship.” Econometrica: Journal of the Econometric Society: 909–938.

Hellmann, Thomas. 2007. “The Role of Patents for Bridging the Science to Market Gap.” Journal of Economic Behavior & Organization 63 (4) (August): 624–647.

Henderson, Rebecca M., Adam B. Jaffe, and Manuel Trajtenberg. 1998. “Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988.” Review of Economics and Statistics 80 (1) (February): 119–127.

Kitch, Edmund W. 1977. “The Nature and Function of the Patent System.” Journal of Law and Economics 20 (2) (October 1): 265–290.

Lacetera, Nicola. 2009a. “Different Missions and Commitment Power in R&D Organizations: Theory and Evidence on Industry-University Alliances.” Organization Science 20 (3) (June): 565 –582. doi:10.1287/orsc.1080.0366.

———. 2009b. “Academic Entrepreneurship.” Managerial & Decision Economics 30 (7) (October): 443–464. doi:10.1002/mde.1461.

Mansfield, Edwin. 1998. “Academic Research and Industrial Innovation: An Update of Empirical Findings.” Research Policy 26 (7-8) (April): 773–776.

Merton, Robert K. 1961. “Singletons and Multiples in Scientific Discovery: A Chapter in the Sociology of Science.” Proceedings of the American Philosophical Society 105 (5) (October 13): 470–486.

———. 1973. The Sociology of Science: Theoretical and Empirical Investigations. University of Chicago Press.
Mokyr, Joel. 2002. The Gifts of Athena: Historical Origins of the Knowledge Economy. Princeton University Press.
Moon, Seongwuk. 2011. "How Does the Management of Research Impact the Disclosure of Knowledge? Evidence from Scientific Publications and Patenting Behavior." Economics of Innovation and New Technology 20 (1): 1–32. doi:10.1080/10438590902750893.
Mowery, David C., Bhaven N. Sampat, and Arvids A. Ziedonis. 2002. "Learning to Patent: Institutional Experience, Learning, and the Characteristics of U. S. University Patents after the Bayh-Dole Act, 1981-1992." Management Science 48 (1) (January 1): 73–89.
Murray, Fiona. 2002. "Innovation as Co-Evolution of Scientific and Technological Networks: Exploring Tissue Engineering." Research Policy 31 (8-9): 1389–1403.
———. 2010. "The Oncomouse That Roared: Hybrid Exchange Strategies as a Source of Distinction at the Boundary of Overlapping Institutions1." The American Journal of Sociology 116 (September): 341–388. doi:10.1086/653599.
Murray, Fiona, Philippe Aghion, Mathias Dewatripont, Julian Kolev, and Scott Stern. 2011. "Of Mice and Academics: Examining the Effect of Openness on Innovation."
Murray, Fiona, and Scott Stern. 2007. "Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge?: An Empirical Test of the Anti-Commons Hypothesis." Journal of Economic Behavior & Organization 63 (4) (August): 648–687.
Nelson, Richard R. 1959. "The Simple Economics of Basic Scientific Research." Journal of Political Economy 67 (3) (June 1): 297–306.
———. 1962. "The Link between Science and Invention: The Case of the Transistor." NBER Chapters: 549–584.
Roach, Michael, and Henry Sauermann. 2010. "A Taste for Science? PhD Scientists' Academic Orientation and Self-Selection into Research Careers in Industry." Research Policy 39 (3) (April): 422–434. doi:10.1016/j.respol.2010.01.004.
Rosenberg, Nathan. 1990. "Why Do Firms Do Basic Research (with Their Own Money)?" Research Policy 19 (2) (April): 165–174. doi:10.1016/0048-7333(90)90046-9.
Rosenberg, Nathan, and Richard R. Nelson. 1994. "American Universities and Technical Advance in Industry." Research Policy 23 (3) (May): 323–348.
Sauermann, Henry, and Paula Stephan. 2013. "Conflicting Logics? A Multidimensional View of Industrial and Academic Science." Organization Science 24 (3) (May 1): 889–909.
Stern, Scott. 1995. "Incentives and Focus in University and Industrial Research: The Case of Synthetic Insulin." In Sources of Medical Technology: Universities and Industry, 157.
———. 2004. "Do Scientists Pay to Be Scientists?" Management Science 50 (6) (June 1): 835–853. doi:10.1287/mnsc.1040.0241.
Stokes, D. E. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Brookings Institution Press.
Thompson, Neil C., David C. Mowery, and Arvids A. Ziedonis. 2013. "Does University Licensing Facilitate or Restrict the Flow of Knowledge and Research Inputs among Scientists?" Accessed November 11.
Williams, Heidi L. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." Journal of Political Economy 121 (1) (February 1): 1–27.

TABLE 1. Variables and definitions

Variable	Definition	Source
Discovery Characteristics		
PAPER-TWIN	Dummy variable for each simultaneous discovery	Twin matching algorithm
SCIENTIFIC POTENTIAL	Yearly rate of publication citations to twin papers by organizations that did not make the discovery	Pubmed
TECHNOLOGICAL POTENTIAL	Yearly rate of patent citations to twin papers by organizations that did not make the discovery	Scopus
YEAR	Year of publication of an article or year of patent filing	WoS; USPTO
Organization Characteristics		
ACADEMIA	Dummy variable equal to 1 if the organization is a university or a government organization; 0 otherwise	Paper itself
PATENT PAIR	Dummy variable equal to 1 if the organization was awarded a patent pair; 0 otherwise	Patent pair matching algorithm
US ORGANIZATION	Dummy variable equal to 1 if the address of organization on the discovery paper is in the US; 0 otherwise	Paper itself
Measures of Knowledge Use		
NPAT	For organization i, NPAT is the count of patents citing the discovery paper that (1) has organization i as assignee, (2) includes one of the discovery author as inventor and (3) was filed after the year of publication	Scopus
NPUB	For organization i NPUS is the count of publications citing the discovery paper that (1) has the discovery organization as address, (2) includes one of the discovery author	Pubmed
PAT_CITES	Count of citations received by patent k in the patent literature by 2013	USPTO
PUB_CITES	Count of citations received by publication k in the publication literature by 2013	Pubmed

TABLE 2. Means and standard deviations

Variable	N	Mean	Std. Dev.	Min	Max
Main Analysis					
# PAPERS PER SIMULTANEOUS DISCOVERY	39	2.31	0.52	2	4
DISCOVERY'S TECHNOLOGICAL POTENTIAL	39	2.37	3.56	0	20.9
DISCOVERY'S SCIENTIFIC POTENTIAL	39	5.94	2.81	0.1	10.8
ACADEMIA	90	0.54	0.50	0	1
US ORGANIZATION	90	0.64	0.48	0	1
PUBLICATION YEAR	90	2000.44	3.65	1994	2008
PATENT PAIR	90	0.24	0.43	0	1
# FOLLOW-ON PATENTS	90	1.90	5.30	0	34
# CITATIONS-WEIGHTED PATENTS	90	17.61	65.30	0	443
# FOLLOW-ON PUBLICATIONS	90	3.56	5.85	0	42
# CITATIONS-WEIGHTED PUBLICATIONS	90	398.13	797.73	0	4132
Quality of Follow-On Patents and Publications					
PATENT BY ACADEMIC ORGANIZATION	171	0.19	0.39	0	1
PATENT BY US ORGANIZATION	171	0.88	0.32	0	1
PATENT APPLICATION YEAR	171	2004.43	4.12	1995	2012
LOG(1+PATENT CITATIONS)	171	1.23	1.28	0	4.1
PUBLICATION BY ACADEMIC ORGANIZATION	320	0.79	0.41	0	1
PUBLICATION BY US ORGANIZATION	320	0.69	0.46	0	1
PUBLICATION YEAR	320	2007.06	4.37	1995	2013
LOG(1+PUBLICATION CITATIONS)	320	2.36	1.33	0	5.7
University-Firm Collaborations (Address Level)					
# ADDRESSES PER PAPER	131	5.90	4.43	2	37
ACADEMIA	773	0.77	0.42	0	1
US ORGANIZATION	773	0.58	0.49	0	1
PATENT PAIR	773	0.06	0.24	0	1
# FOLLOW-ON PATENTS PER ADDRESS	773	0.17	0.74	0	7
# CITATIONS-WEIGHTED FOLLOW-ON PATENTS PER ADDRESS	773	0.30	1.93	0	25
# FOLLOW-ON PUBLICATIONS PER ADDRESS	773	1.43	4.35	0	64
# CITATIONS-WEIGHTED FOLLOW-ON PUBLICATIONS PER ADDRESS	773	42.81	158.98	0	2398

TABLE 3. University-Firm Twins Compared to Other Twins

	UNIVERSITY-UNIVERSITY TWINS	UNIVERSITY-INDUSTRY TWINS	INDUSTRY-INDUSTRY TWINS
# Publications	1146	90	10
# Follow-on Patents	0.2	1.9	3.5
# Citations-Weighted Follow-on Patents	1.5	17.6	58.3
# Follow-on Publications	5.2	3.6	1.3
# Citations-Weighted Follow-on Publications	766.8	398.1	225.8

TABLE 4. University-Firm Twins Compared to Non-Twins

Variable	N	Mean	Std. Dev.	Min	Max
TWIN PAPERS (90 ACROSS UNIVERSITY-INDUSTRY BOUNDARY)					
Yearly Rate of Citation by Patents	90	1.78	2.47	0	15.5
Yearly Rate of Citations by Other Publications	90	52.04	34.74	3.1	147.4
NON-TWINS (180 ADJACENT ARTICLES)					
Yearly Rate of Citation by Patents	180	0.25	0.59	0	4.6
Yearly Rate of Citations by Other Publications	180	17.76	19.28	0.5	131.5

TABLE 5. Impact of Firms and Universities on Follow-On Inventions

	FIXED EFFECT POISSON QML [Incidence-rate ratios in brackets in top line] Estimated coefficients in second line (Robust SEs reported in parentheses)			
	DV = # PATENTS		DV = # CITATION-WEIGHTED PATENTS	
	Marginal impact; no control	Marginal impact w/ controls	Marginal impact; no control	Marginal impact w/ controls
ACADEMIA	[0.186] -1.68 (0.498)***	[0.299] -1.206 (0.438)***	[0.036] -3.311 (0.793)***	[0.053] -2.932 (0.951)***
US ORGANIZATION		[0.616] -0.484 (0.81)		[0.482] -0.73 (1.073)
PATENT PAIR		[4.376] 1.476 (0.871)*		[1.768] 0.57 (1.349)
Observations	50	50	38	38
Log likelihood	-65.18	-57.03	-228.34	-224.34
Paper-twin FE	21	21	16	16

Values are incident rate rations; robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 6. Impact of Firms and Universities on Follow-On Publications

	FIXED EFFECT POISSON QML [Incidence-rate ratios in brackets in top line] Estimated coefficients in second line (Robust SEs reported in parentheses)			
	DV = # PUBLICATIONS		DV = # CITATION-WEIGHTED PUBLICATIONS	
	Marginal impact; no control	Marginal impact w/ controls	Marginal impact; no control	Marginal impact w/ controls
ACADEMIA	[3.103]	[3.049]	[2.506]	[3.078]
	1.132	1.115	0.918	1.124
	(0.247)***	(0.242)***	(0.367)**	(0.321)***
US ORGANIZATION		[1.591]		[0.681]
		0.465		-0.384
		(0.418)		(0.443)
PATENT PAIR		[1.626]		[3.842]
		0.486		1.346
		(0.403)		(0.406)***
Observations	76	76	76	76
Log likelihood	-111.03	-101.35	-2403.52	-2056.93
Paper-twin FE	33	33	33	33

Values are incident rate rations; robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 7. Impact of Universities and Firms on Follow-On Use: Robustness Analysis

	FIXED EFFECT POISSON QML [Incidence-rate ratios in brackets in top line] Estimated coefficients in second line (Robust SEs reported in parentheses)					
	DV = # PATENTS			DV = # PUBLICATIONS		
	Triplets and Quadruplet excluded	Dual affiliations excluded	90% Winsorized DV	Triplets and Quadruplet excluded	Dual affiliations excluded	90% Winsorized DV
ACADEMIA	[0.226]	[0.233]	[0.568]	[3.236]	[4.247]	[2.980]
	-1.487	-1.457	-0.566	1.174	1.446	1.092
	(0.558)***	(0.552)***	(0.312)*	(0.316)***	(0.287)***	(0.245)***
US ORGANIZATION	[0.321]	[0.370]	[1.029]	[0.760]	[0.951]	[1.106]
	-1.138	-0.995	0.0285	-0.275	-0.0498	0.101
	(0.826)	(0.948)	(0.402)	(0.426)	(0.442)	(0.292)
PATENT PAIR	[3.872]	[3.027]	[2.544]	[2.378]	[2.680]	[1.866]
	1.354	1.108	0.934	0.866	0.986	0.624
	(1.301)	(0.930)	(0.532)*	(0.470)*	(0.418)**	(0.355)*
Observations	28	41	50	48	63	76
Log likelihood	-24.70	-46.76	-24.95	-58.66	-66.29	-92.72
Paper-twin FE	14	18	21	24	28	33

Values are incident rate rations; robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 8. Impact of Universities and Firms on Follow-On Use: Interactions

	FIXED EFFECT POISSON QML [Incidence-rate ratios in brackets in top line] Estimated coefficients in second line (Robust SEs reported in parentheses)					
	DV = # PATENTS			DV = # PUBLICATIONS		
ACADEMIA	[0.926]	[0.806]	[0.655]	[3.694]	[1.977]	[2.558]
	-0.0767	-0.216	-0.424	1.307	0.682	0.939
	(0.608)	(1.893)	(0.448)	(0.288)***	(0.502)	(0.306)***
US ORGANIZATION	[1.097]	[1.629]	[1.222]	[1.828]	[1.810]	[1.755]
	0.0927	0.488	0.201	0.603	0.593	0.562
	(0.907)	(1.552)	(0.628)	(0.301)**	(0.323)*	(0.339)*
PATENT PAIR	[0.710]	[0.736]	[1.948]	[1.630]	[1.803]	[1.488]
	-0.342	-0.306	0.667	0.489	0.589	0.397
	(0.402)	(0.482)	(0.483)	(0.309)	(0.349)*	(0.522)
ACADEMIA* TECHNOLOGICAL POTENTIAL	[0.659]			[0.868]		
	-0.416			-0.142		
	(0.207)**			(0.069)**		
ACADEMIA* SCIENTIFIC POTENTIAL		[0.810]			[1.049]	
		-0.21			0.0474	
		(0.333)			(0.069)	
ACADEMIA* PATENT PAPER PAIR			[0.097]			[1.314]
			-2.337			0.273
			(0.812)***			(0.657)
Observations	50	50	50	76	76	76
Log likelihood	-58.26	-61.68	-57.17	-95.63	-99.47	-99.72
Paper-twin FE	21	21	21	33	33	33

Values are incident rate rations; robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 9. The Quality of University and Firm Publications and Patents

	OLS DV = LOG(1+NCITES)					
	NCITES = # PUBLICATION CITATIONS			NCITES = # PATENT CITATIONS		
	Marginal impact; no control; no FE	Marginal impact w/ controls; no FE	Marginal impact w/ controls and FE	Marginal impact; no control; no FE	Marginal impact w/ controls; no FE	Marginal impact w/ controls and FE
ACADEMIA	-0.343	-0.158	-0.0587	-0.904**	-0.894**	0.00889
	(0.22)	(0.14)	(0.21)	(0.33)	(0.33)	(0.50)
YEAR		-0.147***	-0.173***		-0.034	-0.234***
		(0.02)	(0.03)		(0.04)	(0.05)
US ORGANIZATION		-0.0675	-0.478		-0.0471	0.718
		(0.23)	(0.34)		(0.45)	(0.72)
CONSTANT	2.630***	297.2***	349.1***	1.400***	69.51	470.4***
	(0.15)	(48.29)	(55.87)	(0.25)	(84.62)	(94.45)
Observations	320	320	320	171	171	171
R-squared	0.011	0.242	0.242	0.077	0.088	0.367
Paper-twin FE	none	none	33	none	none	21

Robust standard errors in parentheses clustered at the level of the paper-twin
*** p<0.01, ** p<0.05, * p<0.1

TABLE 10. Firm-University Collaboration and Follow-On Inventions

	FIXED EFFECT POISSON QML [Incidence-rate ratios in brackets in top line] Estimated coefficients in second line (Robust SEs reported in parentheses)			
	DV = # PATENTS		DV = # CITATION-WEIGHTED PATENTS	
	Marginal impact; no control	Marginal impact w/ controls	Marginal impact; no control	Marginal impact w/ controls
ACADEMIA	[0.541]	[0.444]	[0.525]	[0.229]
	-0.615	-0.812	-0.645	-1.474
	(0.335)*	(0.406)**	(0.503)	(0.588)**
US AUTHOR		[1.611]		[0.463]
		0.477		-0.77
		(0.238)**		(0.999)
PATENT PAIR		[18.256]		[28.967]
		2.905		3.366
		(0.726)***		(1.201)***
Observations	156	156	87	87
Log likelihood	-106.22	-78.71	-177.17	-115.38
Paper FE	28	28	19	19

Values are incident rate rations; robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 11. Firm-University Collaboration and Follow-On Publications

	FIXED EFFECT POISSON QML [Incidence-rate ratios in brackets in top line] Estimated coefficients in second line (Robust SEs reported in parentheses)			
	DV = # PUBLICATIONS		DV = # CITATION-WEIGHTED PUBLICATIONS	
	Marginal impact; no control	Marginal impact w/ controls	Marginal impact; no control	Marginal impact w/ controls
ACADEMIA	[7.148] 1.967 (0.284)***	[10.571] 2.358 (0.360)***	[4.355] 1.471 (0.277)***	[6.888] 1.93 (0.300)***
US AUTHOR		[3.375] 1.216 (0.366)***		[2.717] 0.999 (0.316)***
PATENT PAIR		[5.620] 1.726 (0.335)***		[8.568] 2.148 (0.429)***
Observations	621	621	613	613
Log likelihood	-855.79	-766.15	-21525.46	-18196.14
Paper FE	105	105	104	104

Values are incident rate rations; robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

FIGURE 1. AN AUTOMATED AND SYSTEMATIC METHOD TO GENERATE A LIST OF SIMULTANEOUS DISCOVERIES (Reproduced from Bikard 2012)

Step 1: Collection of ISI Web of Knowledge data on all research articles from the 15 non-review scientific publications having the highest Journal Impact Factor

(42,106 publications)

↓

Step 2: Using Pubmed and CrossRef, verify the type of article and the complete author list of each of the 1,294,357 references online.

(744,583 unique references)

↓

Step 3: Generation of a database of pairs of all references (a) co-cited at least once, (b) written no more than 1 year apart, (c) having no overlapping author, (d) in which at least 5 citations for each reference are observed in the dataset of citing articles.

(17,050,914 pairs considered; 449,417 pairs selected)

↓

Step 4: Computation of the Jaccard co-citation coefficient for all pairs of references (intersection over the union of forward citations). Highly skewed distribution with a long tale of pairs that are consistently cited in the same papers.

↓

Step 5: Selection of the 2,320 pairs with co-citation coefficient superior to 50% and run a parsing algorithm on all the co-citing articles. Out of these pairs the parsing algorithm could analyze 3 co-citing publications or more in 1,825 cases; 720 pairs have been cited adjacently in 100% of the co-citing articles

FIGURE 2. A SYSTEMATIC METHOD TO FIND PATENT PAIRS TO PAPERS

Step 1: For each author on a discovery paper, collection of all the patents that (a) has their name listed as inventor and (b) was filed the year of paper publication or the year preceding it using the USPTO website.

↓

Step 2: Selection of the patents that emerged for at least two coauthors of a given paper

↓

Step 3: For each organization on a discovery paper, establish whether the organization is an assignee on one of the patents selected in Step 2.

FIGURE 3. CITATIONS TO ACADEMIC AND INDUSTRY TWINS IN SCIENTIFIC JOURNALS

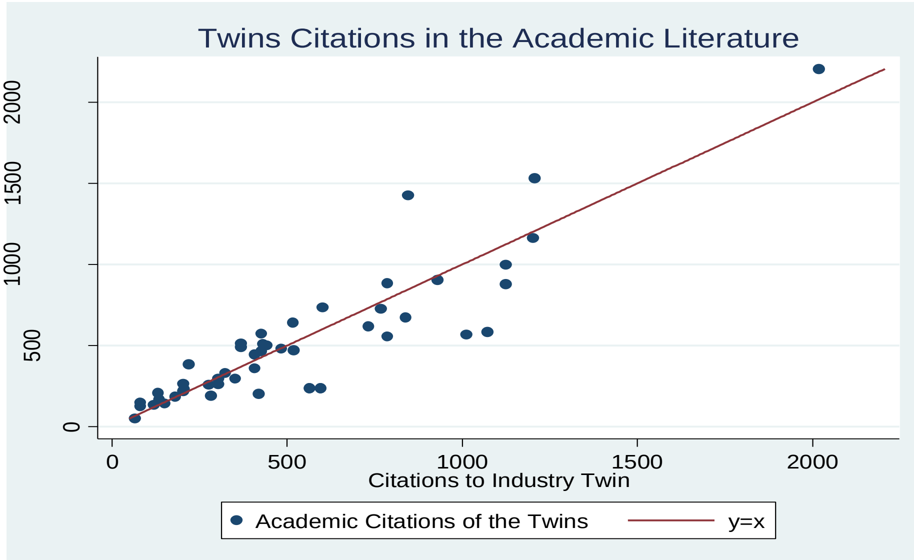


FIGURE 4. OUR SAMPLE IN PERSPECTIVE

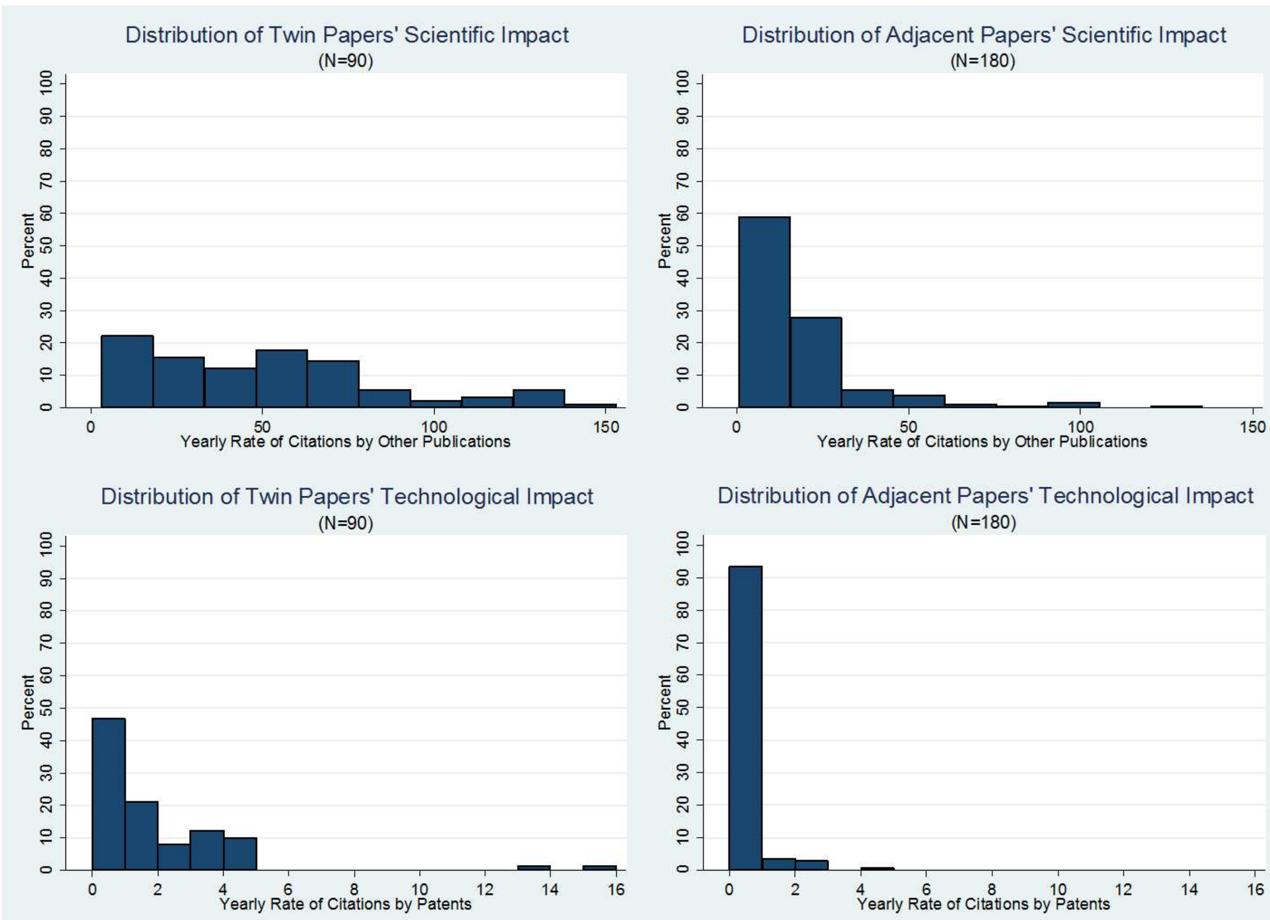
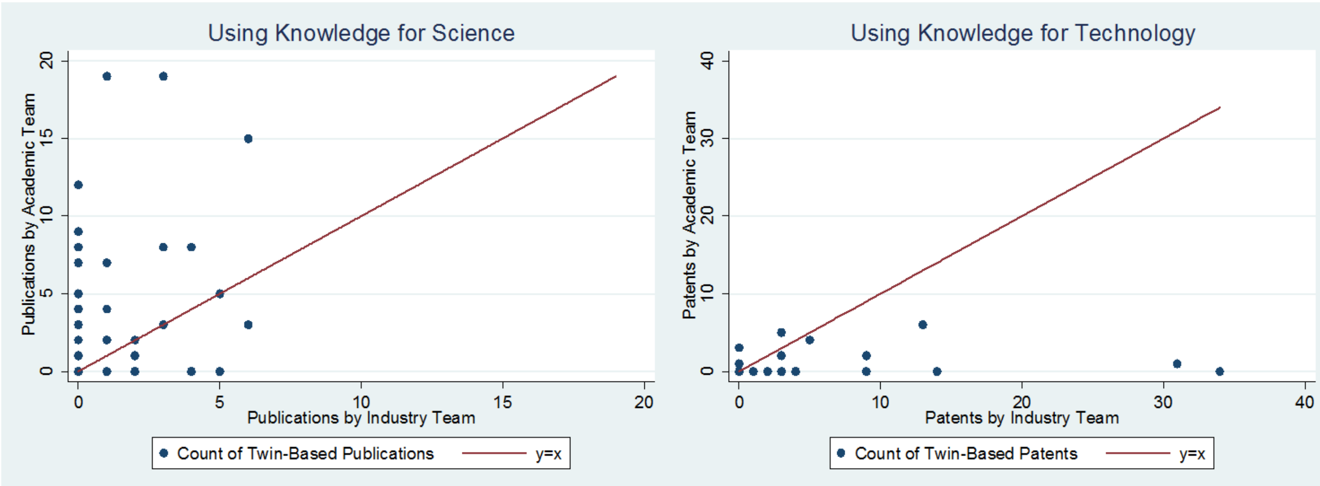


FIGURE 5. MAIN EFFECT: TWO DIFFERENT USES OF SCIENTIFIC KNOWLEDGE*



*Twins involving authors with dual academic and industry affiliations were excluded