London Business School

LBS Research Online

M Bikard, K Vakili and F Teodoridis
When collaboration bridges institutions: the impact of university-industry collaboration on academic productivity
Article

Bikard, M, Vakili, K and Teodoridis, F

(2019)

When collaboration bridges institutions: the impact of university-industry collaboration on academic productivity.

Organization Science, 30 (2). pp. 426-445. ISSN 1047-7039

DOI: https://doi.org/10.1287/orsc.2018.1235

INFORMS
https://pubsonline.informs.org/doi/abs/10.1287/ors...

When Collaboration Bridges Institutions:
The Impact of University–Industry Collaboration on Academic Productivity[1]

Michaël Bikard

London Business School, Strategy and Entrepreneurship Department

Sussex Place, London, NW1 4SA, UK

mbikard@london.edu

Keyvan Vakili

London Business School, Strategy and Entrepreneurship Department

Sussex Place, London, NW1 4SA, UK

kvakili@london.edu

Florenta Teodoridis

University of Southern California, Marshall School of Business

701 Exposition Blvd., Los Angeles, CA 90089, USA

florenta.teodoridis@marshall.usc.edu

Keywords: University-Industry Collaboration; Division of Labor; Scientific Productivity; Innovation; Interorganizational Relationships; Simultaneous Discoveries; Knowledge Production

Acknowledgements: We would like to thank seminar participants at the Wharton Technology & Innovation Conference, Strategy Science Conference, West Coast Research Symposium of Technology Entrepreneurship, and Academy of Management Conference for helpful comments and suggestions. We would also like to thank senior editor Rajshree Agarwal and the anonymous reviewers for their valuable and constructive feedback on the paper. This paper was supported by funding from the London Business School.

[1] All authors contributed equally.

When Collaboration Bridges Institutions:
The Impact of University–Industry Collaboration on Academic Productivity

Abstract

Prior research suggests that academic scientists who collaborate with firms may experience lower publication rates in their collaborative lines of work due to industry's insistence on intellectual property protection through patenting or secrecy. In contrast, we posit that university–industry collaboration can sometimes foster specialization and boost academic contribution to open science. Specifically, research lines with both scientific and commercial potential (i.e., in Pasteur's quadrant) provide an opportunity for a productive division of tasks between academic scientists and their industry counterparts, whereby the former focus on exploiting the scientific opportunities and the latter focus on the commercial ones. The main empirical challenge of examining this proposition is that research projects that involve industry collaborators may be qualitatively different from those that do not. To address this issue, we exploit the occurrence of simultaneous discoveries where multiple scientists make roughly the same discovery around the same time. Following a simultaneous discovery, we compare the follow-on research output of academic scientists who collaborated with industry on the discovery with that of academic scientists who did not. We find that academic scientists with industry collaborators produced more follow-on publications and fewer follow-on patents than did academic scientists without industry collaborators. This effect is particularly salient when the research line has important commercial implications and when the industry partner is an established firm.

Introduction

Frequent collaboration between academia and industry is a hallmark of the knowledge economy. These collaborations are often encouraged at the national level,[2] and many studies have emphasized their strategic importance for firms seeking to gain new knowledge, forge new relationships, and yield higher R&D productivity (Cockburn and Henderson 1998; Zucker, Darby, and Armstrong 2002; Owen-Smith and Powell 2004; Lacetera 2009). However, the opposite impact, that of industry collaboration on academic science, remains less understood and more contentious (e.g., Murray 2010; Perkmann et al. 2013). The academic institutional environment highly differs from that of industry (Merton 1973; Dasgupta and David 1994), and some scholars have raised concerns that the corporate emphasis on commercialization might

[2] See http://www.nih.gov/news-events/news-releases/nih-fund-collaborations-industry-identify-new-uses-existing-compounds and http://www.nih.gov/news-events/news-releases/nih-launches-collaborative-program-industry-researchers-spur-therapeutic-development as examples of funding schemes to boost collaboration between academia and industry.

erode academic contribution to open science and weaken the norms of science in academia (Blumenthal, Campbell, et al. 1996; Mowery et al. 2001; Perkmann and Walsh 2009).

At its core, this concern rests on the view that industry's institutional environment conflicts with that of academia. Specifically, firms' emphasis on commercialization and appropriation of their intellectual property rights through patenting and secrecy can limit their academic collaborators' contribution to open science (Blumenthal, Campbell, et al. 1996; Louis et al. 2001; Goldfarb 2008; Perkmann and Walsh 2009; Toole and Czarnitzki 2010; Murray 2010; Evans 2010b; Shibayama, Walsh, and Baba 2012; Czarnitzki, Grimpe, and Toole 2014), induce a guarded behavior among scientists (Blumenthal, Campbell, et al. 1996; Campbell et al. 2000, 2002), and eventually reduce public disclosure of research output in terms of scientific publications (Czarnitzki, Grimpe, and Toole 2014; Lee 2000; Thursby and Thursby 2002). The adverse effects of industry collaboration on academic scientists' productivity and output could be particularly heightened in areas with higher commercial potential and thus higher appropriation risks. The problem is compounded by concerns of shifts in academic institutional norms from freedom of science (Merton 1973) toward knowledge transfer through commercialization (e.g., Azoulay, Ding, and Stuart 2009). Prior research shows that industrial practices of secrecy, patenting, and entrepreneurship have already spread across the halls of academia to some extent (e.g., Argyres and Liebeskind 1998; Stuart and Ding 2006; Bercovitz and Feldman 2008).

In contrast, we propose that, under certain circumstances, collaboration might provide an opportunity for productive specialization across institutional environments. Rather than channeling industry practices into academia, industry collaboration might instead reinforce traditional characteristics of the academic institutional environment by increasing publication and lowering patenting rates. In research lines that have both scientific and commercial potential, collaboration between academia and industry might lead to a more productive division of tasks and responsibilities, where each side would be able to focus on what their institutional environment rewards most: academia on more fundamental scientific insights, and industry on commercialization (Merton 1973; Aghion, Dewatripont, and Stein 2008; Sauermann and Stephan 2013).

Prior research suggests that academic institutions possess a relative competitive advantage in investigating more basic, fundamental questions, whereas firms maintain a relative advantage in development and commercialization. Scientists sort into either academia or industry according to their preferences. Scientists with a stronger "taste for science" or preferences for non-pecuniary returns are more likely to join academia; scientists with relatively stronger preference for applied research and monetary incentives sort into industry (Roach and Sauermann 2010; Agarwal and Ohyama 2013). In addition, the academic institutional environment is particularly amenable to the exchange of knowledge and materials, as well as to curiosity-driven research, whereas firms' resources and expertise foster efficient commercialization (Bush 1945; Nelson 1959; Merton 1973; Rosenberg and Nelson 1994; Aghion, Dewatripont, and Stein 2008; Sauermann and Stephan 2013). Therefore, from firms' point of view, collaborating with academia enables them to focus on areas in which they are the strongest, while relying on their academic partners to explore the projects' more fundamental aspects. Active collaboration during a project can ensure that the knowledge created by the academic side will quickly and seamlessly transfer to the corporate partner. Likewise, collaborating with industry can give academic scientists access to resources, skills, and equipment that might be valuable for fulfilling the scientific potential of a line of research, without the pressure to spend much time on its commercialization. In a sense, then, in research lines that have both scientific and commercial potential, collaboration might be an arena in which the specificities of the academic institutional environment can be leveraged and reinforced.

Hence, by offering the opportunity for specialization and a more productive allocation of tasks on lines of research with both scientific and commercial potential, industry collaboration might bolster academic contribution to open science. We expect academic scientists who collaborate with industry in projects that offer both scientific and commercial potential to experience higher subsequent publication rates within the collaborative lines of works. We further expect to see increased levels of specialization when academic scientists collaborate with established companies, rather than startups, given that the former usually possess better commercialization capabilities and generally more resources to leverage during the collaboration.

The empirical challenge in testing these arguments is considerable because scientists do not randomly choose their research projects. Firms and academic scientists strategically select the research questions they want to pursue. Thus, the type of research projects pursued through university–industry collaboration is likely to be systematically different from those solely pursued in academia. This makes comparisons between academic productivity and engagement with commercialization across different types of projects ill-suited for testing our predictions. Since academic scientists are on average more likely to work on applied projects when they collaborate with firms than when they do not, a finding that academic researchers are on average less scientifically productive when they work with firms might simply indicate that academic scientists work with firms on lines of research that are on average less scientifically promising.

To address this challenge, our empirical strategy holds the line of research constant by exploiting a phenomenon known as multiple or simultaneous discoveries (Merton 1957). In our data, simultaneous discoveries are instantiated as scientific achievements that are published at the same time by two or more different scientific teams. When two teams make roughly the same discovery at about the same time, they face the same level of scientific and commercialization opportunities following that discovery.

We use a sample of 33 such simultaneous discoveries, each involving 2 or 3 teams. In each of those instances, at least one academic team collaborated with a firm, and at least one other did not. For each team disclosing a simultaneous discovery, we track the corresponding author's follow-on contribution to open science and engagement with commercialization for the line of research associated with the simultaneous discovery. We use scientific publications as indicators of contribution to open science and patents as indicators of engagement with the commercialization process.[3] Specifically, we count the corresponding authors' publications and patents that were built upon the simultaneous discovery – i.e., follow-on publications and follow-on patents respectively.

[3] We consider patenting of scientific discoveries as a step in the long commercialization process of discoveries.

Consistent with our predictions, our results suggest that following a simultaneous discovery, academic scientists with industry collaborators produce more follow-on publications and fewer follow-on patents than academic scientists who made the same discovery without an industry collaborator. In our data, academic scientists who collaborate with firms produce 50% more citation-weighted publications. Our results also suggest that scientists who collaborate with firms produce fewer citation-weighted patents than their academic peers who do not collaborate with industry. Moreover, we find that the results are driven by discoveries that combine both high scientific and high commercial potential (as opposed to high scientific potential only) and discoveries where the industry partner is an established firm (as opposed to a startup). We further show that the higher levels of publication output built upon discoveries with industry partners do not come at the expense of a decline in publication output outside the collaborative research line.

Our empirical design based on simultaneous discoveries ensures that the estimated differences in research output are not driven by differences in the types of research projects pursued by scientists who choose to collaborate with industry. However, the teams in our sample are not randomly formed. Hence, differences in publication output might be attributed to several potential sources. It is possible that scientists who choose to collaborate with industry have higher publication and lower patenting capabilities in the first place (Mindruta 2013). It is also possible that the estimated differences might be driven by physical and monetary capital provided by industry collaborators, independent of their active collaboration in the project. Finally, consistent with our theory, differences could be driven by academic scientists' greater ability to specialize in knowledge production when they collaborate with firms. To distinguish between these factors, we show that, before publishing their simultaneous discoveries, academic scientists with industry collaborators had no more (and in fact slightly fewer) publications than scientists without industry collaborators. We further show that our results are not driven by simple industry sponsorship in the form of financial or physical capital. Overall, our findings are consistent with the idea that in research lines with both scientific and commercial potential, collaboration can provide an arena in which the complementarities between academia and industry enhance the research productivity of academic scientists.

Our paper makes three main contributions. First, the current literature on cross-institutional collaboration has highlighted that those partnerships can constitute a bridge or a channel across which travels not only knowledge (Zucker, Darby, and Armstrong 2002; Owen-Smith and Powell 2004) but also norms and practices (Czarnitzki, Grimpe, and Toole 2014; Evans 2010b). We complement this work by showing that cross-institutional collaboration constitutes also an opportunity for productive specialization in creative work. Second, our results contribute to the literature that has highlighted the potentially negative impact of tight academia–industry relationships for the academic institutional environment (Argyres and Liebeskind 1998; Shibayama, Walsh, and Baba 2012). The traditional argument is that increased closeness between academia and industry fosters hybridization and decreases the distinctiveness of the academic institutional environment. As a result, academic contribution to open science might suffer. In contrast, we argue that, when scientific knowledge can contribute to both fundamental understanding and the development of commercial technologies, increased closeness between academia and industry can *reinforce* the distinctive features of the academic institutional environment. Finally, empirical studies of the impact of scientific collaboration for creative performance are complicated by the fact that collaboration is a choice (e.g., Mindruta 2013). Here we propose that some of the issues associated with this selection process can be addressed by focusing on simultaneous discoveries as sets of "natural experiments."

Industry Collaboration and Academic Productivity

Collaboration as a Conduit through Which Industrial Practices Flow into Academia

Historically, the key mandate of universities has been to generate and disseminate knowledge (Bush 1945). A defining characteristic of academic researchers is autonomy, such that scientists are free to select the projects they undertake and have creative control over the research methods they pursue and how they choose to disseminate their findings (Aghion, Dewatripont, and Stein 2008). With the increased collaboration between industry and academia over the past few decades, many have expressed concerns that too much involvement with industry's commercialization activities and profit-seeking agendas may distort the institutional norms that have long governed academic research (Mowery et al. 2001). In particular, scholars have argued that collaboration between industry and academia acts as a bridge between the two institutional environments, and hence, industrial orientation toward economic gains and secrecy might travel over that bridge and erode the traditional academic emphasis on openness and publication.

Scholars discussing the negative implications of university–industry collaboration point out that industry collaboration is likely to harm academic productivity because industrial practices of appropriation through secrecy and intellectual property rights clash with academic incentives to publish and disseminate the results of their research fast and widely (Merton 1973; Dasgupta and David 1994; Stephan 1996). Firms often impose publication restrictions on their academic collaborators within the collaborative lines of research (e.g., Lee 2000; Thursby and Thursby 2002). Surveys of academic researchers have associated collaboration with industry with higher levels of secrecy, less sharing in general, and more publication delays (Blumenthal, Causino, et al. 1996; Blumenthal et al. 1997; Louis et al. 2001; Evans 2010b; Shibayama, Walsh, and Baba 2012; Czarnitzki, Grimpe, and Toole 2014). For example, in a survey of 210 life-science companies, Blumenthal et al. (1996) found that among the companies that supported academic research, 82% asked academic scientists to keep their findings confidential for some time to allow for patent applications, and that 56% require confidentiality for longer than the time needed for patent application. Relationships with industry might deter sharing because of the emphasis on confidentiality, but the impact might also be deeper, as sharing norms might progressively shift from generalized to more direct forms of exchange (Shibayama, Walsh, and Baba 2012). Similar problems were reported in cases of industry sponsorship: an indirect form of university–industry collaboration. For example, in a recent survey of 1,060 German academic researchers, Czarnitzki, Grimpe, and Toole (2014) found that, compared with other forms of funding, industry sponsorship increases the expected probability of publication delay from 14% to 33%, and that of secrecy (complete or partial ban on publishing research results) from 11.2% to 35%. Other research shows that patenting a scientific discovery—as is often the case in university–industry collaborations—may negatively affect scientists' investment in follow-on research based on that discovery (Murray and Stern 2007; Williams 2013). Overall, then, academic scientists working with industry are likely to find it on average more difficult to publish their research results.

In addition to the issue of publication restrictions, industry collaboration could also harm academic productivity by inducing scientists to undertake commercialization-related activities that are unlikely to lead to scientific publications (Goldfarb 2008; Perkmann and Walsh 2009; Toole and Czarnitzki 2010). Some empirical studies have linked academic scientists' involvement with industry to a decrease in their contribution to open science. In particular, Toole and Czarnitzki (2010) study the impact on academic productivity of life-scientists' involvement in for-profit firms through the SBIR program. Using a case-cohort sampling design, they find that the yearly publication rate of academic scientists decreases by 19% after they become involved in the program, which indicates a considerable "brain drain" away from academic science and toward commercial activities.

The potential negative impact of industry collaboration on academic scientists' contribution to open science seems especially alarming considering that the scientists who work with firms tend to be some of the most productive in their fields (Blumenthal, Campbell, et al. 1996; Zucker and Darby 1996; Stuart and Ding 2006; Toole and Czarnitzki 2010; Mindruta 2012).

Collaboration as an Arena for a Productive Division of Tasks between Academia and Industry

In studying collaborations across academia and industry, prior research has generally considered (a) that collaboration acts as a bridge and (b) that this bridge erodes the differences opposing each institutional environment. In so doing, that literature has therefore uncovered an important negative facet of cross-institutional collaboration.

In contrast, we argue that under certain circumstances, academia–industry collaboration may foster specialization across institutional environments. Collaboration, we argue, is not just a channel through which practices and norms can diffuse; it is also an arena in which the complementarities across institutional environments can be leveraged. In projects with both scientific (basic) and commercial (applied) opportunities, academia and industry can mutually benefit if each party focuses on what it does best; i.e., if academia focuses on developing fundamental understanding, and industry focuses on commercialization. In such projects, industry collaboration might reinforce the particularities of the academic institutional environment by creating an opportunity for academic scientists to specialize and devote their efforts mostly toward exploiting the scientific potential of the project while being minimally distracted by the commercialization process.

A large stream of research argues that firms can benefit from collaborating with academia on research projects (Cockburn and Henderson 1998; Zucker, Darby, and Armstrong 2002; Furman and MacGarvie 2009). Academic publications can improve R&D efficiency if that scientific knowledge is relevant to the firm (Nelson 1982; Fleming and Sorenson 2004; Bikard 2018). Academic findings can also open up new market opportunities for firms (Zucker, Darby, and Brewer 1998) and help them overcome technological barriers (Cohen, Nelson, and Walsh 2002).

We argue that academic scientists can likewise benefit from collaborating with industry. The presence of an industry partner can help academic scientists achieve greater levels of specialization by focusing on their basic research activities and leave the commercialization aspects, such as patenting and licensing, to their industry collaborators. Past research suggests that academia is relatively less efficient than industry when it comes to commercialization (Rosenberg and Nelson 1994; Aghion, Dewatripont, and Stein 2008). To the extent that academic scientists also care about the ultimate application of their research (Murray 2010; D'Este and Perkmann 2011; Stephan 2012), industry collaboration might also increase their motivation to invest in those lines of research. Besides, industry collaboration might even help scientists invest more heavily in lines of research with commercialization opportunities. Prior research suggests that industry involvement in a research domain, particularly in the form of patenting, may dissuade academic scientists from investing in follow-on research due to potential litigation risks (Heller and Eisenberg 1998; Murray and Stern 2007). The under-exploitation of such research lines could give firms incentives to correct the resulting deficit through collaboration.

The benefits from greater specialization are likely to be further amplified by the fact that academic scientists might find industry to be a valuable source of skills, equipment, and material and financial resources, but also of ideas (Mansfield 1995; D'Este and Perkmann 2011; Tartari and Breschi 2012). D'Este and Perkmann (2011), for example, surveyed 1,088 grant holders from the UK's Engineering and Physical Sciences Research Council who had engaged with industry and asked them about their motivation for doing so. Interestingly, their results highlight four key drivers: access to funding, commercialization, learning, and access to in-kind resources. Industry and academia usually differ in how they approach research; they may use different perspectives to formulate research questions, and different methods to investigate them (Rosenberg 1994; Mansfield 1995; Siegel et al. 1999; Evans 2010a). Hence, collaboration with industry can give academic scientists access to new and different insights and knowledge assets, which could in turn increase their research output. A long stream of research on creativity suggests that by bringing together diverse perspective, insights, and knowledge, collaboration can help innovative teams break away from intellectual lock-in and explore new, otherwise unexplored territories (Hargadon and Sutton 1997; Reagans, Zuckerman, and McEvily 2004; Uzzi and Spiro 2005; Fleming, Mingo, and Chen 2007; Teodoridis 2017). Using a theoretical model, Hong and Page (2004) demonstrate how a diverse team of randomly selected individuals may outperform a homogeneous team of best-performing ones because individuals with varied backgrounds and perspectives could evaluate the same opportunities differently, and hence exploit a more heterogeneous set of opportunities.

Following these arguments, we postulate that scientists collaborating with industry on certain projects might generate higher levels of follow-on contribution to open science (e.g., publication) and lower involvement in commercialization activities (e.g., patenting). Industry collaboration in a particular line of research might accentuate some institutional characteristics that are distinctively academic. We therefore predict the following:

> *Hypothesis 1 (H1): For a given research project with both academic and applied potential, an academic scientist with an industry collaborator is likely to produce more follow-on scientific publications than an academic scientist without an industry collaborator.*
>
> *Hypothesis 2 (H2): For a given research project with both academic and applied potential, an academic scientist with an industry collaborator is likely to produce fewer follow-on patents than an academic scientist without an industry collaborator.*

An efficient and productive specialization and division of tasks and responsibilities between academia and industry along their domains of expertise requires that the collaborative project at hand offer high potential for both scientific research and commercialization, which enables each party to focus on what it does best. To offer additional insight, we focus on the moderating effect of the commercial potential of a line of research.[4]

The above predictions challenge the idea that the pursuit of economic returns in industry negatively affects the research output of collaborating academic scientists. If, as prior research suggests, appropriation concerns lead to more secrecy, one should expect secrecy and intellectual property protection to play a more significant role in projects with more commercial potential. In contrast, our arguments suggest that higher commercial potential may provide better conditions for specialization through a productive allocation of tasks and efforts. Hence, we predict the following:

Hypothesis 3 (H3): Commercial potential of a research project positively moderates the relationship between academia–industry collaboration and the follow-on scientific publications of academic scientists.

Furthermore, we expect the experience level of industry partners to positively moderate the benefits that academic scientists accrue from their collaboration with industry. Established companies, as opposed to startups, have deeper industry-specific knowledge, better commercial capabilities, and generally more resources to share with their academic partners. Given the central role of these factors in facilitating specialization and a productive division of tasks and responsibilities between the two types of institutions, we expect that academic scientists who collaborate with established companies yield more benefit than those who collaborate with startups. This prediction also contradicts what prior research suggests. Established companies are usually more embedded in the industry mindset of rent-seeking and protecting intellectual property rights, and hence are expected to be more cautious with publishing the outcomes of their research projects. Startups, on the other hand, may use these publications as signals of their quality and be more supportive of contribution to open science. However, we argue that the knowledge assets and

[4] While we believe scientific potential as a moderating factor could also offer interesting deeper insights, we are limited by our empirical context, in which all research projects have considerable scientific potential. The lowest number of academic citations to a paper in our sample is 66, which already puts that paper in the top 1% of highly cited papers in its respective field.

commercialization capabilities of established companies better support the hypothesized productive collaboration between academia and industry. Hence, we predict the following:

Hypothesis 4 (H4): Collaboration with established companies positively moderates the relationship between academia–industry collaboration and the follow-on scientific publications of academic scientists.

Data and Empirical Strategy

The core challenge in examining the impact of industry–academia collaboration on subsequent research productivity of academic scientists is that scientists strategically select the research questions they want to pursue. Thus, the research projects that are pursued through university–industry collaboration may fundamentally differ from those pursued solely in academia or industry. Consequently, any estimated effect of university–industry collaboration on subsequent research productivity captures both the effect of collaboration and the effect of choosing a particular research question.

In an ideal setting, we would observe the same research question assigned randomly and simultaneously to two teams: one composed only of academic collaborators and one that includes industry collaborators. We would then want to observe follow-on publication and patenting efforts of academic scientists from both teams within the respective line of research and to measure the differences between those efforts. In the absence of such an ideal experiment, we use a novel empirical approach exploiting the existence of simultaneous discoveries.

Simultaneous discoveries are scientific discoveries that are independently conducted by two or more teams of scientists and reported in what we call paper twins: papers published on the same discovery around the same time. Paper twins are hence dual instances of the same scientific discovery by different teams in different environments. The following example resulted from a discovery that was simultaneously made by scientists from the Harvard Medical School and scientists at the University of Utah and Heidelberg University working in collaboration with Myriad Pharmaceutical:

Strack et al. (September 2003) "AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding" *Cell*

(Affiliations: Department of Cancer Immunology and AIDS Dana-Farber Cancer Institute and Department of Pathology, Harvard Medical School)

"HIV-1 and other retroviruses exit infected cells by budding from the plasma membrane, a process requiring membrane fission. The primary late assembly (L) domain in the p6 region of HIV-1 Gag mediates the detachment of the virion by recruiting host Tsg101, a component of the class E vacuolar protein sorting (Vps) machinery. We now show that HIV Gag p6 contains a second region involved in L domain function that binds AIP1, a homolog of the yeast class E Vps protein Bro1. Further, AIP1 interacts with Tsg101 and homologs of a subunit of the yeast class E Vps protein complex ESCRT-III. AIP1 also binds to the L domain in EIAV p9, and this binding correlates perfectly with L domain function. These observations identify AIP1 as a component of the viral budding machinery, which serves to link a distinct region in the L domain of HIV-1 p6 and EIAV p9 to ESCRT-III."

von Schwedler et al. (September 2003) "The protein network of HIV budding" *Cell*

(Affiliations: Department of Biochemistry, University of Utah; Department of Pathology, Heidelberg University; Myriad Pharmaceuticals Inc.)

"HIV release requires Tsg101, a cellular factor that sorts proteins into vesicles that bud into multivesicular bodies (MVB). To test whether other proteins involved in MVB biogenesis (the class E proteins) also participate in HIV release, we identified 22 candidate human class E proteins. These proteins were connected into a coherent network by 43 different protein-protein interactions, with AIP1 playing a key role in linking complexes that act early (Tsg101/ESCRT-I) and late (CHMP4/ESCRT-III) in the pathway. AIP1 also binds the HIV-1 p6 and EIAV p9 proteins, indicating that it can function directly in virus budding. Human class E proteins were found in HIV-1 particles, and dominant-negative mutants of late-acting human class E proteins arrested HIV-1 budding through plasma and endosomal membranes. These studies define a protein network required for human MVB biogenesis and indicate that the entire network participate in the release of HIV and probably many other viruses."

These excerpts describe two sets of independent findings regarding the role of the same protein components in the HIV viral budding process. One of the papers (Strack et al. 2003) was written by a research team within academia, the other (von Schwedler et al. 2003) in a collaborative effort between scientists in academia and a firm (Myriad Pharmaceuticals Inc.). The two papers were published back-to-back in the September 2003 volume of the journal *Cell*.

The use of simultaneous discoveries helps us control for the strategic selection into particular research questions by different scientists and hence isolate the effect of university–industry collaboration. In particular, we assume that when two different teams report the same discovery, they face the same level of follow-on scientific and commercial opportunity to pursue based on that discovery. This approach allows us to observe academic scientists' behavior with respect to subsequent knowledge production based on the discovery when influenced by industry collaboration, while controlling for the otherwise hard-to-observe counterfactual of the same project undertaken absent industry collaboration.

In addition, we assume that the collaborative teams on simultaneous discoveries are formed in expectation of certain complementarities. Indeed, in addition to strategically choosing their research projects, scientists also choose their collaborators. However, there might be other factors determining a choice of collaborators that do not necessarily speak to complementarities as assumed by our theory. For example, scientists who collaborate with industry may have higher publishing and lower patenting capabilities. In our empirical analysis, we address these concerns directly, while discussing limitations.

Data

Our dataset contains 72 scientific publications disclosing 33 simultaneous discoveries.[5] As noted above, in each paper twin in our sample, one paper is authored by only academic scientists, and the other results from collaboration between academia and industry. The main (corresponding) author on all publications is an academic scientist.[6]

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 within the same parentheses, or adjacently (Cozzens 1989). Figure 1A in the

[5] Twenty-seven of the simultaneous discoveries were reported in two publications, and 6 were disclosed in three papers.

[6] The corresponding author is clearly identified on each publication as the main point of contact between the authorship team and all external parties. The corresponding author is generally the more senior individual who takes responsibility for the research, provides the intellectual input, and approves the research design and protocols. In about 70% of the papers in our sample, the corresponding author is the same as the last author. In most cases where the corresponding author is the first author, we could confirm that the first author was the senior scientist running the lab. Our focus on the corresponding author as the main author is in line with previous literature (e.g., Furman and Stern 2011; Furman, Jensen, and Murray 2012; Zucker et al. 1998).

appendix summarizes the algorithm used to construct the sample of simultaneous discoveries. The algorithm is based on the method developed by Bikard (2012).

We collected data about each paper from three bibliographical databases: ISI Web of Science, Scopus, and PubMed. Details about the main (corresponding) scientists were collected from various public sources, such as the scientists' public profiles. Publication data (through October 2013) were collected from two bibliographical databases: Scopus and PubMed. Most discoveries in our sample are in life sciences in areas such as cell biology, stem cell, and immunology. The sample includes discoveries in physics as well.

Our patent dataset is the result of the intersection of two sets of patent data which we built separately. First, we collected all the patents that build on each simultaneous discovery. To do so, we built a script that used the Harvard Dataverse patent data (Li et al. 2014), which is based on USPTO data, to identify the set of patents that cite either of the papers disclosing the simultaneous discovery. Patent citations to papers indicate that the inventors drew on the knowledge disclosed in the scientific paper as part of the invention process (e.g., Azoulay, Graff Zivin, and Sampat 2012; Roach and Cohen 2012). Therefore, we consider those patents to constitute inventive activity building on the simultaneous discovery. Second, we collected all the U.S. patents awarded by each corresponding author in our sample. To do so, we searched the USPTO website for each scientist's name. To address the issue of different inventors having the same name (disambiguation), we assessed each patent separately and established the inventor's identity based on application date, inventor's address, and patent topic. Our dataset of follow-on patents by the academic scientists is the intersection between these two sets.

Empirical Strategy

For each simultaneous discovery, we compare the follow-on research output of the main authors after the publication of the simultaneous discovery. We define follow-on research output of a scientist as the count of scientific papers that list the scientist as an author and cite her initial paper reporting the simultaneous discovery. By fixing the initial discovery, we can attribute differences in the follow-on research output of scientists to the composition of their teams for their initial discovery, i.e., the difference in their involvement with industry. To better isolate the effect of university–industry collaboration, we further control for several observable characteristics of the main scientists and their respective teams. Table 1 lists the set of variables used in our analyses and explains how they are defined and constructed.

-- Table 1 approximately here –

Formally, we use the following regression equation to estimate the effect of university–industry collaboration on follow-on publications of corresponding scientists on each twin paper:

$$follow\_on\_publication\_count_{ip} = f(\beta academia\_industry\_collaboration_{ip} + \alpha X_{ip} + discovery_p + \varepsilon_{ip})$$

where $follow\_on\_publication\_count_{ip}$ is the number of academic publications of scientist i citing the initial simultaneous discovery p after its publication. X_{ip} is a vector of controls at the scientist-discovery level, including scientist's academic experience, her past academic publication and patent stock, a dummy capturing whether the scientist was affiliated with an institution in the United States, the number of authors listed on the paper reporting the initial simultaneous discovery p, the share of authors affiliated with institutions in the United States, the share of female authors, and a dummy capturing whether the paper was supported by any industry grants or not. Simultaneous discovery fixed effects are captured by $discovery_p$. The inclusion of discovery fixed effects ensures that all the comparisons are between the main scientists who made a simultaneous discovery around the same time and hence faced the same subsequent scientific and commercial opportunity. In other words, we do not rely on comparisons across discoveries and only compare scientists who made the same discovery around the same time. Because the dependent variable is a count, we use a Poisson model with robust standard errors. The results are robust to the use of a more restricted set of control variables as well as to alternative specifications including negative binomial and ordinary least squares (OLS) estimation models.

Following our theoretical arguments, to the extent that collaboration between academia and industry leads to a better allocation of tasks and responsibilities, and the associated positive effect on research output of academic scientists overcomes the potential negative effects proposed in prior literature, we expect β to be positive and significant. Otherwise, in the absence of any positive effect, or where the negative effects argued in prior literature are greater than the positive effects argued in this study, β would be negative or non-significant.

We use the same regression method to identify the impact of academia–industry collaboration on follow-on patenting by each scientist (H2). The follow-on patenting efforts are measured as the count of patents that cite the original discovery (in their non-patent references) and list the main scientist behind the discovery as an inventor. Note that the follow-on patents are not necessarily patents granted on the original discovery itself. Rather, these are generally patents filed on subsequent commercial developments informed by the simultaneous discovery. Following our theoretical arguments, we expect the effect of university–industry collaboration to be negative.

To test H3 and H4, we add interaction terms between our dummy variable $academia\_industry\_collaboration_{ip}$ and variables capturing whether the discovery has high commercial potential or not (H3) and whether the industry partner on discovery p was an established company or not (H4). We use the total count of patents that have cited any of the twin papers reporting a simultaneous discovery as a proxy for the commercial potential of that discovery. We do not include the variable alone in the regressions since its effect is captured by discovery fixed effects. We then categorize discoveries that are above the median in commercial potential as discoveries with high commercial potential. The results are robust to using a continuous measure of commercial potential. To identify whether a company partner was established or not, we use the age of the company partner at the time of each discovery. Companies with more than 10 years of experience at the time of collaboration are categorized as established. Again, we only include the variable in its interaction form because its direct effect is captured by the discovery fixed effect. The results are robust to neighboring cutoff years for this categorization. We also perform several additional robustness analyses that we discuss in more detail below.

Boundary Conditions

Note that our theory and empirical analysis are not generalizable to the whole spectrum of scientific projects but instead apply to those projects with both scientific and commercial potential. Hence, if one thinks of the landscape of scientific publications in terms of Donald Stokes' (1997) typology, our twin papers would be better representatives of papers in Pasteur's quadrant, where discoveries have both scientific and commercial potential. This selection corresponds directly to our theoretical argument.

How large is the population of papers to which our findings are potentially generalizable? While it is difficult to assess how many papers, of the massive universe of scientific publications, have both scientific and commercial potential, we can note that papers that stand at the interface between science and technology are known to be particularly important (Ahmadpoor and Jones 2017). We can also provide some perspective by comparing our sample of paper twins to other (non-twin) manuscripts published by the same authors. To do this, we collected all publications by the corresponding authors in our sample that were published in the same year as the twin papers. We used the same publication year as the twins to ensure that the difference between the twin and non-twin publications by each scientist is not driven by her tenure or changes in the opportunity landscape in her field over time. A main author in our sample has produced on average 9.1 other papers (st.dev=8.6) in the same year as her twin paper. Next, we categorized all papers published by the corresponding authors into four categories: 1) twin papers without industry collaborators; 2) twin papers with industry collaborators; 3) non-twin, academia-only papers by scientists who had no industry collaborators on their twin papers; and 4) non-twin, academia-only papers by scientists who had industry collaborators on their twin papers. We compared these four categories on two main dimensions: the count of academic citations, and whether the papers were cited in any subsequent patent (0/1 dummy). We used citations in subsequent patents as a proxy for the commercial potential of a scientific publication. We further log-normalized the number of academic citations to address their skewed distribution. Figures 1 and 2 compare the distributions of these two variables across the four categories. The twin papers have on average higher rates of both paper citations and patent citations. This is expected considering the nature of twin discoveries. The distribution of paper citations as well as the existence of a patent citation is tilted toward the right in both figures. However, there is a good overlap between the distributions of twin papers and non-twin, academia-only papers. Specifically, Figure 3 overlays the distributions for twins without industry collaborators (first category) onto the distributions for academia-only, non-twins (second and fourth categories). The overlap suggests that twin papers still represent an important part of academia-only publications.

-- Figures 1, 2 and 3 approximately here –

Descriptive Statistics

We provide summary statistics for our main variables in Table 2. The average number of follow-on publications for each author-paper was 20.9, and the average number of follow-on patents was 0.44. Approximately 22% of corresponding authors obtained a follow-on patent citing their twin papers in the sample. The oldest of the simultaneous discoveries dated back to 1996, and the most recent occurred in 2008. On average, about 10 authors collaborated on each discovery paper. In 64% of discovery papers, the corresponding author was affiliated with a U.S. institution. Also, 61% of the authors on each discovery paper were affiliated with U.S. institutions. A typical main author had about 16 years of experience and about 113 papers and 2 patents before the publication of the discovery paper. Approximately 30% of authors on the discovery papers were women, and about 26% of the discovery papers were supported by industry grants.

-- Table 2 approximately here –

The list of all universities and non-for-profit research organizations involved with the discoveries in our sample is relatively long. Of the 750 scientists involved in these simultaneous discoveries, about 77 researchers are affiliated with a company and the rest with an academic institution. In total, there are 183 unique academic or non-for-profit research institutions in our sample. Some of the most frequent institutions are the Institute for Cancer Research (London, UK), Harvard University, the University of California, the University of Michigan, Yale University, the Salk Institute, the University of Texas, Dana-Farber Cancer Institute, Max Planck Institute, the University of Iowa, the University of Utah, Columbia University, Duke University, Imperial College London, Scripps Research Institute, and Rockefeller University.

Furthermore, in total, 25 unique companies are involved as collaborators on the papers in our sample. A few companies are involved with more than one discovery in our sample. We define startup firms as firms that were 10 years old or younger at the time of their collaboration on the simultaneous discovery.

The established companies in our sample are Novartis, Merck, Schering-Plough, Bayer, GSK, Genentech, Xerox Corporation, the EMMES Corporation, Mitsubishi, Elan Corporation, and Amgen. Of these, the youngest is Amgen, which was established in 1980. The list of startups includes Millennium Pharmaceuticals, NeuroSpheres Ltd., Athersys, Virologic, Caprion Pharmaceuticals, Ligand Pharmaceuticals, Lexicon Pharmaceuticals, Regeneron Pharmaceuticals, Perlegen Sciences, Sangamo Biosciences, ATABIS GmbH, deCODE Genetics, and Myriad Pharmaceuticals. The startups in our sample were all founded after 1991.

Results

As a first step, we report some simple comparative statistics. In approximately 59% of the cases, the academic authors who had industry collaborators produced more follow-on research than those who had no industry collaborators. Also, when comparing the twins, after excluding the cases where none of the corresponding authors of either twin produced any follow-on patents, in 64% of the cases the academic scientists who had an industry collaborator produced fewer follow-on patents than the academic scientists who had no industry collaborators on their twin discovery. Table 3 further compares the normalized number of follow-on and pre-discovery papers by authors with and without industry collaborators. Due to the count nature of data and the substantial across-twin variance in our sample, we need to normalize the publication rates to be able to meaningfully compare the two groups of scientists. We use the within-twin averages for normalization. For example, take two academic scientists as main authors on two papers reporting a simultaneous discovery, who have 40 and 50 pre-discovery papers and 10 and 20 follow-on papers, respectively. To normalize the numbers, we divide the number of pre-discovery papers of each scientist by 45 (the average of 40 and 50) and the number of follow-on papers by 15 (the average of 10 and 20). The normalized figures show that academic scientists with industry collaborators produce approximately 15% fewer pre-discovery papers than those without industry collaborators. In contrast, the number of follow-on papers by scientists with industry collaborators is about 8% larger than that of scientists without industry collaborators. The difference between the change in the number of pre-discovery and follow-on papers by scientists with and without industry collaborators is significant at the 10% level. These comparative statistics are in line with H1. However, they cannot effectively distinguish between cross-twin and within-twin variances. We address this issue in our estimations by using twin fixed effects in all specifications.

-- Table 3 approximately here –

In Table 4, we report results of our main estimating equation: the impact of university–industry collaboration on the count of main scientists' follow-on academic publications and patents. Model 4-1 shows the impact of university–industry collaboration on the count of follow-on academic publications. Consistent with H1, the results suggest that scientists with industry collaborators produce 37% more follow-on publications than scientists without industry collaborators. The difference amounts to approximately 7.7 additional follow-on papers published based on the original discovery—an increase of 0.44 standard deviations. In model 4-2, we adjust our dependent variable to take into account the quality of follow-on academic publications. As in prior research, we use the number of citations that an academic publication received as a proxy for its quality. The results show that industry collaboration has a similar positive effect on the quality-adjusted number of follow-on academic publications. Industry collaboration is associated with an additional 1,542 citation-weighted follow-on publications—an increase of 0.35 standard deviations.[7]

-- Table 4 approximately here --

Model 4-3 shows the estimated impact of university–industry collaboration on follow-on patenting by the main scientists. The estimates suggest that, compared with those having no industry collaborators, scientists who collaborated with industry partners on their simultaneous discoveries produce more than 99% fewer patents building on their discovery. In model 4-4, we again adjust the number of patents based on their quality, as measured by the number of citations. The estimated effect of industry collaboration is similar to that reported in model 4-3. The results are consistent with H2. The large estimates suggest that the effect is largely driven by whether the corresponding authors patent or not. Hence, in Table 5 we repeated our estimations using Logit and Probit estimation models with a dummy dependent variable

[7] All subsequent analyses include the full set of controls; however, due to space considerations, we do not report these values. Tables with the full set of controls are available upon request.

indicating whether the main author has any follow-on patents (DV=1) or not (DV=0) based on her simultaneous discovery.[8] The results are qualitatively similar to those reported in Table 4; however, the effects are not significant at the 10% level (P=0.13 in the Probit model and P=0.16 in the Logit model). The standard deviations are larger potentially due to the loss of approximately 20% of the observable variance from the cases where both corresponding authors of matched twin papers have produced follow-on patents based on their simultaneous discovery. Overall, our finding on the impact of industry collaboration on follow-on patenting persists. Nevertheless, given our small sample size and the rare nature of patenting incidents, the estimations are rather sensitive to the choice of statistical model. This implies that the interpretation of magnitude of patenting effects should be considered with care. While we provide these figures, we do so with the goal of drawing attention to the theoretically important effect on patenting. We hope that future research can more robustly estimate the magnitude of this effect.[9]

-- Table 5 approximately here --

In model 6-1 of Table 6, we investigate how the commercial potential of a discovery moderates the effect of industry collaboration on scientists' follow-on research output. The estimate indicates that the positive effect of academia–industry collaboration on scientists' research productivity is driven by the subset of discoveries with high commercial potential. For the category of discoveries with high commercial potential, scientists with industry collaborators produce 54% more follow-on publications than scientists without industry collaborators, calculated based on the sum of the direct effect and the interaction effect. The estimated effect is equivalent to 11 additional follow-on publications—an increase of approximately

[8] We could not obtain convergence on our estimations based on Logit and Probit models with the full set of control variables due to the limited observable variance and few degrees of freedom. Hence, in our robustness checks with Logit and Probit estimations, we use a more limited set of control variables. More specifically, we only control for whether the main author is affiliated with an institution in the U.S. (particularly because we use the sample of patents filed with the USPTO) and the main author's past patenting experience.

[9] To provide additional evidence on patenting behavior, we note that 13 firms in our sample have obtained patents based on the simultaneous discoveries on which they had collaborated with academia. To put it differently, in 40% of the cases firms in our sample obtained patents that cited their simultaneous discoveries. In the absence of an appropriate counterfactual, we cannot accurately evaluate whether this number is large or small. However, a recent analysis by Ahmadpoor and Jones (2017) suggest that approximately 20% of papers produced by firms are subsequently cited in patents. This suggests that collaboration with academia in our sample may have boosted firms' follow-on commercialization outputs. It is also possible that for discoveries published later in our sample, we do not observe enough time periods to better evaluate the output of the commercialization efforts of collaborating firms.

0.65 standard deviations. In contrast, for discoveries with low commercial potential, the effect is negative and not statistically significant. This negative effect suggests that academia–industry collaboration on projects with low commercial potential may lead to negative effects in research productivity of academic scientists, as argued in prior research. In other words, in the absence of the conditions for specialization and a productive division of tasks and responsibilities, collaboration with industry may potentially harm academic scientists' research output. Model 6-2 repeats the same estimation using the quality-adjusted number of follow-on publications as the dependent variable. The estimates are similar and overall consistent with our prediction in H3.

-- Table 6 approximately here –

In models 6-3 and 6-4 of Table 6 we present estimation results testing our fourth hypothesis, namely the moderating role of collaboration with established companies versus startups. Consistent with our prediction, the estimated interaction effect in model 6-3 indicates that scientists with industry collaborators produced approximately 67% more follow-on publications (equivalent to approximately 14 additional follow-on publications or an increase of 0.8 standard deviations) than scientists without industry collaborators. In contrast, there is no significant difference between scientists who collaborated with startups (companies with 10 years or less of experience) and those without industry collaborators. Model 6-4 repeats the previous estimation with a quality-adjusted number of follow-on publications as the dependent variable. The results are consistent and support H4.

In Table 7 we repeat the estimations reported in interaction models 6-2 and 6-4 using negative binomial and OLS regressions with discovery fixed effects. The results are similar to those reported with the Poisson model. While the estimated negative effect of industry collaboration on projects with low commercial potential is larger in these estimates, the net positive effect on citation-weighted follow-on publications based on discoveries with high potential (calculated as the sum of the direct effect and the interaction effect) is comparable to that reported in Table 6.

-- Table 7 approximately here –

Alternative Explanations and Robustness Tests

While we control for each scientist's past publication stock in all models reported in Table 4, one may still be concerned that the positive impact of academia–industry collaboration on scientists' follow-on research output may be driven by the selection of more productive scientists into collaboration with industry. In other words, it is possible that scientists with higher research productivity are more likely to collaborate with industry, and that the same scientists continue to publish more based on the simultaneous discoveries. To address this concern, we test the relationship between the likelihood of collaboration with industry and each author's past research productivity, patenting activity, academic experience, and location. Since the dependent variable is a 0/1 dummy variable, we use a logit estimation model. We also include discovery fixed effects to ensure that the comparison takes place between scientists reporting the same discovery. The estimates in model 8-1 of Table 8 suggest that collaboration with industry is negatively associated with past research productivity in our dataset. Further, the estimates indicate that collaboration with industry is not correlated with past patenting activity of scientists or with their academic experience. In model 8-2, we repeat the estimations using an OLS regression and find similar results. Overall, the estimates suggest that the positive effects reported in Table 4 (in support of H1) are not driven by selection of more productive scientists into collaboration with industry.

-- Table 8 approximately here --

For each pair of corresponding authors on twin papers, we also compare their non-twin, academia-only papers published in the same year as their twin papers. We compare the non-twin papers on three main dimensions: the number of academic citations received, whether they received any patent citations subsequently, and the number of authors on each paper. A higher similarity in non-twin papers by the corresponding authors behind a simultaneous discovery provides further evidence that the differences in follow-on research output of scientists who collaborated with industry are not driven by scientists' average higher quality. The results are presented in Table A1 in the appendix. The estimates show no significant difference in the non-twin papers produced by the authors involved in simultaneous discoveries—whether they had an industry partner or not. Overall, the results support the assumption that the output of corresponding authors who had industry collaborators at the time of publishing their twin papers was similar to the output of those who did not have collaborators from industry.

Even though the previous analysis shows that the differences in the experience of the corresponding authors on twin papers is unlikely to explain the differences in their follow-on research output, it is still possible that the corresponding authors with industry collaborators had a more senior or more experienced team of co-authors on their side. In other words, the increase in their follow-on research output may be associated with the presence of more experienced co-authors and not that of an industry partner. To address this concern, we first identified the status of every author on every twin paper at the time of its publication using publicly available information. We then categorized the authors into research faculty, university research staff (such as graduate students and postdocs), industry scientists, and unidentified. We could identify the status of 60% of the authors of twin papers in our sample. Models 1 to 3 of Table A2 in the appendix show the results of estimating the difference in number of faculty, the number of research staff (grad students and postdocs), and the number of individuals with unidentified status between the teams with industry collaborators and the teams without. Furthermore, we collected all pre-discovery publications by all authors in each authorship team using the Scopus database. To the extent that the team with industry collaboration has more experienced individuals, we should expect that team to also have a higher level of total pre-discovery publications. Model 4 of Table A2 shows the results of estimating the authorship team's total number of pre-discovery publications on whether the team involved industry partners. We find no significant difference between the two teams.

Another possible alternative interpretation of our results is that the increase in scientists' follow-on research output linked to industry collaboration may have come at the expense of research output in areas outside the discovery's line of research. In other words, what we are interpreting as an increase in research output driven by specialization and a complementary allocation of tasks may simply be due to a shift in the research direction of scientists collaborating with industry. In Table 9, we test the impact of academia–industry collaboration on scientists' research output outside the discovery line of research. The effect on both the simple count and the quality-adjusted count of publications, reported in models (9-1) and (9-2) respectively, shows that we find no evidence that industry collaboration affects scientists' research output in areas that are not directly related to the collaborative line of research.

-- Table 9 approximately here –

Furthermore, to ensure that our results are driven not by industry sponsorship but by active collaboration with industry, we also test the effect of industry grants on follow-on research productivity of scientists using a larger sample of 1,151 papers written by academic scientists and disclosing 540 simultaneous discoveries. Some of those academic scientists received industry sponsorship for their discovery, and others did not. To build the sample of industry grants, we manually collected the acknowledgments sections of all the papers in the sample and extracted all mentions of funding support from industry. Our results in Table 10 suggests that industry sponsorship has no significant effect on follow-on research productivity of scientists, and that the effects reported here rely on the mechanism of active collaboration between academia and industry, as hypothesized.

-- Table 10 approximately here –

Our results are also robust to the inclusion of a more restricted set of control variables (to allow for more degrees of freedom) as shown in Table A3 in the appendix.

Discussion and Conclusion

Our paper draws upon literatures on collaboration, institutions, and creativity to examine the implications of industry–academia collaborations for academic research. Specifically, we model the influence of industry collaboration on academic scientists' follow-on publishing and patenting rates. This question has received increasing attention in light of widespread concerns that industry collaboration might corrupt traditional academic norms. In this study we argue that when a research line has both scientific and commercial potential, the collaboration between academia and industry can enable specialization and a better allocation of tasks and responsibilities across the two types of institutions, leading to greater complementarity. More specifically, academic scientists can achieve higher levels of scientific output by gaining access to skills, insights, resources, and equipment from firms, while leaving the applied and commercial aspects of the research line to their industry partner.

We face an important empirical challenge in investigating the impact of industry collaboration for academic scientists because those scientists do not randomly sort into research projects that involve industry collaboration. Prior findings that industry collaboration is associated with fewer publications and more patents in academia might be a consequence of norm contamination, but they might also reflect the fact that academic scientists collaborate with firms on their more applied projects with lower scientific potential. This paper attempts to address the challenge of project selection by focusing on a set of simultaneous discoveries, in which an academic scientist collaborating only with other academic scientists makes the same discovery as one collaborating with industry. We observe 33 such events, which we use as a small set of natural experiments to investigate the existence of systematic differences in the subsequent publishing and patenting patterns of scientists with and without industry collaborators for those lines of work.

Holding the line of research constant, we find that academic scientists who collaborate with industry produce significantly more follow-on publications and fewer patents than academic scientists who do not. Collaboration with industry is associated with 50% more follow-on citation-weighted publications and 99% fewer follow-on citation-weighted patents. As mentioned above, the magnitude effect of patenting should be interpreted with care considering our small sample size and the rarity of patenting events in the sample. Moreover, the effects do not appear to be driven either by the sorting of more capable scientists into collaborations with industry, or by the fact that firms might induce academic scientists to invest more in their line of research of interest at the expense of other research projects. Rather, we find evidence consistent with specialization and the existence of a better allocation of tasks and responsibilities between academia and industry, wherein academic scientists focus more on publishing and invest less in patenting. In line with this argument, our main effects are most pronounced for the lines of research that have high commercial potential and those projects that involve collaborations with mature, established companies. To be clear, our findings rely on the assumption that academic and industry scientists engage in collaboration based on some foreseen complementarities. In other words, our results do not suggest that for any research project in Pasteur's quadrant, any random collaboration between academia and industry would have positive benefits for academic scientists. The complementarities need to be there for these collaborations to be fruitful.

Our conversations with several authors of simultaneous discoveries in our sample further corroborate our quantitative estimations. For example, when asked about the role of university and industry in their collaborative projects on drug development, a leading faculty member at a prominent West Coast university noted:

> More and more, you are seeing the enlightened companies getting their molecules out there into academic labs so people can play around with them. [Companies] make drugs and reagents, but they often don't have the very sophisticated and elegant systems that we have to interrogate things in vivo…We have these fairly elegant, genetically-engineered models that really recapitulate what happens in human cancer.

Another academic scientist from a prominent university made a similar remark, highlighting that industry collaboration can provide access to capabilities that can speed up research:

> The fact that we collaborated with [firm] meant that we could go faster. That did give us a competitive advantage because they had this capability of doing high throughput two-hybrid screening.

Scientists in industry shared similar sentiments regarding the complementarities between academia and industry in their collaborative projects. As an example, when asked about his simultaneous discovery and the allocation of tasks and efforts between his team in industry and his collaborators in academia, a senior scientist in a large East Coast pharmaceutical company responded:

> [In] this case we were sharing our mice with the academic labs… [and] they would do the analysis of the mice. The company would set up a material transfer agreement that handles the legalities of sending the mice over to the academic labs. There would be hypotheses that would be set up front in terms of the experiments that would be carried on in the academic lab. Then we'll send the mice, they would do the experimental stuff and then they looked to see if they find something interesting.

> The company cannot do all the science itself. So one of the things I did at [my company] was to make reagents and send them to the academic labs as soon as possible, so you maximize the value of the reagents that you are generating. So it's a way to connect with academic labs, a way to know the science before it comes out as a publication, it's a way of helping the field move forward.

When asked whether the publication of findings would help the company's competitors, he responded that

> …making new medicine has a 99% failure rate. So if you're going to be super-conservative and be closed and siloed away from everyone, you are actually increasing the odds that you are not going to be able to deliver anything. So I think what you gain is more than what you lose by being transparent.

Overall, our results emphasize a hitherto under-recognized positive aspect of industry collaboration for academic productivity. By conceptualizing university–industry collaboration as a channel through which norms travel, prior literature has highlighted that the misalignment in objectives between academic and industrial scientists often leads to tensions among collaborators. These tensions, in turn, might affect creative performance negatively (Blumenthal, Campbell, et al. 1996; Louis et al. 2001; Perkmann and Walsh 2009; Toole and Czarnitzki 2010; Evans 2010). We propose that this view of cross-institutional collaboration is incomplete. Under some circumstances, collaboration can be a source of productive specialization (Bikard, Murray, and Gans 2015; Vakili 2016; Teodoridis 2017; Teodoridis, Bikard, and Vakili 2018). When collaboration bridges institutions, the differences in the skills and objectives of individuals across the institutional boundary can open the door to an efficient distribution of tasks and responsibilities, potentially leading to net gains in productivity.

Scientists in industry and in academia have different approaches to scientific research, but they also tend to work on different types of projects (Sauermann and Stephan 2013). While practices such as secrecy and patenting are well-suited to efficient commercialization, they are barriers to sharing and publication when it comes to fundamental research (Arrow 1962; Aghion, Dewatripont, and Stein 2008). Thus, industry relationships that foster the spread of commercial practices in universities could hurt fundamental inquiry in academia (Campbell et al. 2000; Evans 2010b; Murray and Stern 2007; Murray 2010; Shibayama, Walsh, and Baba 2012). One should note, however, that fundamental inquiry and commercialization are not two ends of the same continuum. Research projects can have at the same time high scientific and high commercial potential (Stokes 1997), therefore raising the question of the appropriate institutional environment for this type of projects. Our results suggest that in those cases, the institutional environments of industry and academia are in some ways complements rather than substitutes. For those projects, proximity to industry does not erode the distinctive features of the academic institutional environments but appears instead to reinforce them.

Our findings are not without limitations and therefore open the door to further research. Our empirical strategy relies on using simultaneous discoveries to ensure that the compared scientists are working on the same line of research. Still, industry collaboration is not randomly assigned in our data; hence it is reasonable to assume that there may still be some systematic differences between the scientists who work with firms and those who do not. As mentioned, our theory assumes that academic and industry scientists engage in collaboration based on some foreseen complementarities. While we do not specifically identify the source of such complementarities, we take steps to control for some observable characteristics that may speak more to individual team-member capabilities rather than team complementarities (Agarwal and Ohyama 2013). Future research should consider the role of additional attributes (e.g., location, proximity, skill complementarities) to determine the optimal conditions under which collaboration with industry on projects with both scientific and commercialization potential is most efficient for knowledge production in academia.

Institutional diversity creates opportunities for specialization and a more productive organization of creative work. When the same work presents differential values across disparate institutional environments, collaboration can make it possible to take advantage of cross-institutional differences. This type of cross-institutional collaboration can be found in various other forms of public–private partnerships (e.g., Battilana and Lee 2014). Our study calls for future research on contingencies and factors under which, instead of creating unproductive conflicts of interests, misallocation of resources and efforts, and wasteful organizational resistance, such institutional diversity can increase the productivity of all involved parties.

References

Ahmadpoor, Mohammad and Benjamin F. Jones, 2017. "The dual frontier: Patented inventions and prior scientific advance." *Science*, 357(6351): 583-587.

Agarwal, Rajshree, and Atsushi Ohyama. 2013. "Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists." *Management Science* 59 (4): 950–70. https://doi.org/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): 617–35.

Ahmadpoor, Mohammad, and Benjamin F. Jones. 2017. "The Dual Frontier: Patented Inventions and Prior Scientific Advance." *Science* 357 (6351): 583–87. https://doi.org/10.1126/science.aam9527.

Argyres, Nicholas S., and Julia Porter Liebeskind. 1998. "Privatizing the Intellectual Commons: Universities and the Commercialization of Biotechnology." *Journal of Economic Behavior & Organization* 35 (4): 427–54. https://doi.org/10.1016/S0167-2681(98)00049-3.

Arrow, K. 1962. "Economic Welfare and the Allocation of Resources for Invention." In *R.R. Nelson (Ed.), The Rate and Direction of Inventive Activity*, 609–25. Princeton, NJ: Princeton University Press.

Azoulay, Pierre, Waverly W. Ding, and Toby Stuart. 2009. "The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research Output." *The Journal of Industrial Economics* 57 (4): 637--676.

Battilana, Julie, and Matthew Lee. 2014. "Advancing Research on Hybrid Organizing – Insights from the Study of Social Enterprises." *Academy of Management Annals* 8 (1): 397–441. https://doi.org/10.1080/19416520.2014.893615.

Bercovitz, Janet, and Maryann Feldman. 2008. "Academic Entrepreneurs: Organizational Change at the Individual Level." *Organization Science* 19 (January): 69–89. https://doi.org/10.1287/orsc.1070.0295.

Bikard, Michaël. 2018. "Made in Academia: The Effect of Institutional Origin on Inventors' Attention to Science." *Organization Science*. Forthcoming.

Bikard, Michaël, Simultaneous Discoveries as a Research Tool: Method and Promise (September 30, 2012). Available at SSRN: https://ssrn.com/abstract=2329605

Bikard, Michaël, Fiona Murray, and Joshua S. Gans. 2015. "Exploring Trade-Offs in the Organization of Scientific Work: Collaboration and Scientific Reward." *Management Science* 61 (7): 1473–95. https://doi.org/10.1287/mnsc.2014.2052.

Blumenthal, David, Eric G. Campbell, Anderson MS, Causino N, and Louis K. 1997. "Withholding Research Results in Academic Life Science: Evidence from a National Survey of Faculty." *JAMA* 277 (15): 1224–28. https://doi.org/10.1001/jama.1997.03540390054035.

Blumenthal, David, Eric G. Campbell, Nancyanne Causino, and Karen Seashore Louis. 1996. "Participation of Life-Science Faculty in Research Relationships with Industry." *N Engl J Med* 335 (23): 1734–39. https://doi.org/10.1056/NEJM199612053352305.

Blumenthal, David, Nancyanne Causino, Eric G. Campbell, and Karen Seashore Louis. 1996. "Relationships between Academic Institutions and Industry in the Life Sciences — An Industry Survey." *New England Journal of Medicine* 334 (6): 368–74. https://doi.org/10.1056/NEJM199602083340606.

Bush, Vannevar. 1945. "Science: The Endless Frontier." *Transactions of the Kansas Academy of Science (1903-)* 48 (3): 231–64.

Campbell, Eric G., B. R. Clarridge, M. Gokhale, L. Birenbaum, S. Hilgartner, N. A. Holtzman, and D. Blumenthal. 2002. "Data Withholding in Academic Genetics: Evidence From a National Survey." *JAMA* 287 (4): 473.

Campbell, Eric G, Joel S Weissman, Nancyanne Causino, and David Blumenthal. 2000. "Data Withholding in Academic Medicine: Characteristics of Faculty Denied Access to Research

Results and Biomaterials." *Research Policy* 29 (2): 303–12. https://doi.org/10.1016/S0048-7333(99)00068-2.

Cockburn, Iain, and Rebecca Henderson. 1998. "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery." *The Journal of Industrial Economics* 46 (2): 157–82.

Cohen, Wesley, Richard Nelson, and John 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 University of New York Press.

Czarnitzki, Dirk, Christoph Grimpe, and Andrew A. Toole. 2014. "Delay and Secrecy: Does Industry Sponsorship Jeopardize Disclosure of Academic Research?" *Industrial and Corporate Change*, May, dtu011. https://doi.org/10.1093/icc/dtu011.

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

D'Este, Pablo, and Markus Perkmann. 2011. "Why Do Academics Engage with Industry? The Entrepreneurial University and Individual Motivations." *The Journal of Technology Transfer* 36 (3): 316–39. https://doi.org/10.1007/s10961-010-9153-z.

Evans, James A. 2010a. "Industry Induces Academic Science to Know Less about More." *The American Journal of Sociology* 116 (2): 389–452.

———. 2010b. "Industry Collaboration, Scientific Sharing, and the Dissemination of Knowledge." *Social Studies of Science* 40 (5): 757–91. https://doi.org/10.1177/0306312710379931.

Fleming, Lee, Santiago Mingo, and David Chen. 2007. "Collaborative Brokerage, Generative Creativity, and Creative Success." *Administrative Science Quarterly* 52 (3): 443–75.

Fleming, Lee, and Olav Sorenson. 2004. "Science as a Map in Technological Search." *Strategic Management Journal* 25: 909–28.

Furman, Jeffrey, and Megan MacGarvie. 2009. "Academic Collaboration and Organizational Innovation: The Development of Research Capabilities in the US Pharmaceutical Industry, 1927–1946*." *Industrial and Corporate Change* 18 (5): 929–61. https://doi.org/10.1093/icc/dtp035.

Goldfarb, Brent. 2008. "The Effect of Government Contracting on Academic Research: Does the Source of Funding Affect Scientific Output?" *Research Policy* 37 (1): 41–58. https://doi.org/10.1016/j.respol.2007.07.011.

Hargadon, Andrew, and Robert Sutton. 1997. "Technology Brokering and Innovation in a Product Development Firm." *Administrative Science Quarterly* 42 (4): 716–49. https://doi.org/10.2307/2393655.

Heller, Michael A., and Rebecca S. Eisenberg. 1998. "Can Patents Deter Innovation? The Anticommons in Biomedical Research." *Science* 280 (5364): 698–701. https://doi.org/10.1126/science.280.5364.698.

Hong, Lu, and Scott E. Page. 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences of the United States of America* 101 (46): 16385–89. https://doi.org/10.1073/pnas.0403723101.

Lacetera, Nicola. 2009. "Different Missions and Commitment Power in R&D Organizations: Theory and Evidence on Industry-University Alliances." *Organization Science* 20 (3): 565–82. https://doi.org/10.1287/orsc.1080.0366.

Lee, Yong S. 2000. "The Sustainability of University-Industry Research Collaboration: An Empirical Assessment." *The Journal of Technology Transfer* 25 (2): 111–33. https://doi.org/10.1023/A:1007895322042.

Louis, Karen Seashore, Lisa M. Jones, Melissa S. Anderson, David Blumenthal, and Eric G. Campbell. 2001. "Entrepreneurship, Secrecy, and Productivity: A Comparison of Clinical and Non-Clinical Life Sciences Faculty." *The Journal of Technology Transfer* 26 (3): 233–45. https://doi.org/10.1023/A:1011106006976.

Mansfield, Edwin. 1995. “Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing.” *The Review of Economics and Statistics* 77 (1): 55–65.
Merton, Robert K. 1957. “Priorities in Scientific Discovery: A Chapter in the Sociology of Science.” *American Sociological Review* 22 (6): 635–59.
———. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.
Mindruta, Denisa. 2013. “Value Creation in University-Firm Research Collaborations: A Matching Approach.” *Strategic Management Journal* 34 (6): 644–65. https://doi.org/10.1002/smj.2036.
Mowery, David C., Richard R. Nelson, Bhaven N. Sampat, and Arvids A. Ziedonis. 2001. “The Growth of Patenting and Licensing by U.S. Universities: An Assessment of the Effects of the Bayh-Dole Act of 1980.” *Research Policy* 30 (1): 99–119. https://doi.org/10.1016/S0048-7333(99)00100-6.
Murray, Fiona. 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): 341–88. https://doi.org/10.1086/653599.
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): 648–87. https://doi.org/10.1016/j.jebo.2006.05.017.
Nelson, Richard. 1959. “The Simple Economics of Basic Scientific Research.” *Journal of Political Economy* 67 (3): 297–306.
———. 1982. “The Role of Knowledge in R&D Efficiency.” *The Quarterly Journal of Economics*, The Quarterly Journal of Economics, 97 (3): 453–70.
Owen-Smith, Jason, and Walter W. Powell. 2004. “Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community.” *Organization Science* 15 (1): 5–21.
Perkmann, Markus, Valentina Tartari, Maureen McKelvey, Erkko Autio, Anders Broström, Pablo D’Este, Riccardo Fini, et al. 2013. “Academic Engagement and Commercialisation: A Review of the Literature on University–industry Relations.” *Research Policy* 42 (2): 423–42.
Perkmann, Markus, and Kathryn Walsh. 2009. “The Two Faces of Collaboration: Impacts of University-Industry Relations on Public Research.” *Industrial and Corporate Change* 18 (6): 1033–106.
Reagans, Ray, Ezra Zuckerman, and Bill McEvily. 2004. “How to Make the Team: Social Networks vs. Demography as Criteria for Designing Effective Teams.” *Administrative Science Quarterly* 49 (1): 101–33.
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): 422–34. https://doi.org/10.1016/j.respol.2010.01.004.
Rosenberg, Nathan. 1994. *Exploring the Black Box: Technology, Economics, and History*. Cambridge University Press.
Rosenberg, Nathan, and Richard Nelson. 1994. “American Universities and Technical Advance in Industry.” *Research Policy* 23 (3): 323–48. https://doi.org/10.1016/0048-7333(94)90042-6.
Sauermann, Henry, and Paula Stephan. 2013. “Conflicting Logics? A Multidimensional View of Industrial and Academic Science.” *Organization Science* 24 (3): 889–909. https://doi.org/10.1287/orsc.1120.0769.
Shibayama, Sotaro, John P. Walsh, and Yasunori Baba. 2012. “Academic Entrepreneurship and Exchange of Scientific Resources Material Transfer in Life and Materials Sciences in Japanese Universities.” *American Sociological Review*, August, 0003122412452874. https://doi.org/10.1177/0003122412452874.
Siegel, Donald, David Waldman, Jonathan Silberman, Albert Link, C. A. Santa Rosa, and A. S. U. West. 1999. “Assessing the Impact of Organizational Practices on the Performance of University Technology Transfer Offices: Quantitative and Qualitative Evidence.”
Stephan, Paula. 1996. “The Economics of Science.” *Journal of Economic Literature* 34 (3): 1199–1235.
———. 2012. *How Economics Shapes Science*. Cambridge, Massachusetts: Harvard University Press.

Stokes, Donald E. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Brookings Institution Press.
Stuart, Toby E., and Waverly W. Ding. 2006. "When Do Scientists Become Entrepreneurs? The Social Structural Antecedents of Commercial Activity in the Academic Life Sciences." *American Journal of Sociology* 112 (1): 97–144. https://doi.org/10.1086/502691.
Tartari, Valentina, and Stefano Breschi. 2012. "Set Them Free: Scientists' Evaluations of the Benefits and Costs of University–industry Research Collaboration." *Industrial and Corporate Change* 21 (5): 1117–47. https://doi.org/10.1093/icc/dts004.
Teodoridis, Florenta. 2017. "Understanding Team Knowledge Production: The Interrelated Roles of Technology and Expertise." *Management Science*, July. https://doi.org/10.1287/mnsc.2017.2789.
Teodoridis, Florenta, Michaël Bikard, and Keyvan Vakili. 2018. "The Pace of Change and Creative Performance: Specialist and Generalist Mathematicians at the Fall of the Soviet Union." SSRN Scholarly Paper ID 2911141. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2911141.
Thursby, Jerry G., and Marie C. Thursby. 2002. "Who Is Selling the Ivory Tower? Sources of Growth in University Licensing." *Management Science* 48 (1): 90–104.
Toole, Andrew A., and Dirk Czarnitzki. 2010. "Commercializing Science: Is There a University 'Brain Drain' from Academic Entrepreneurship?" *Management Science* 56 (9): 1599–1614. https://doi.org/10.1287/mnsc.1100.1192.
Uzzi, Brian, and Jarrett Spiro. 2005. "Collaboration and Creativity: The Small World Problem." *American Journal of Sociology* 111 (2): 447–504.
Vakili, Keyvan. "Collaborative promotion of technology standards and the impact on innovation, industry structure, and organizational capabilities: Evidence from modern patent pools." *Organization Science* 27, no. 6 (2016): 1504-1524.
Williams, Heidi L. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." *Journal of Political Economy* 121 (1): 1–27. https://doi.org/10.1086/669706.
Zucker, Lynne G., and Michael R. Darby. 1996. "Star Scientists and Institutional Transformation: Patterns of Invention and Innovation in the Formation of the Biotechnology Industry." *Proceedings of the National Academy of Sciences* 93 (23): 12709–16.
Zucker, Lynne G., Michael R. Darby, and Jeff S. Armstrong. 2002. "Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology." *Management Science* 48 (1): 138–53.
Zucker, Lynne G., Michael R. Darby, and Marilynn B. Brewer. 1998. "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises." *The American Economic Review* 88 (1): 290–306.

Figure 1. Academic Citation Distribution of Twin and Non-Twin Papers Published by Corresponding Authors in the Same Year

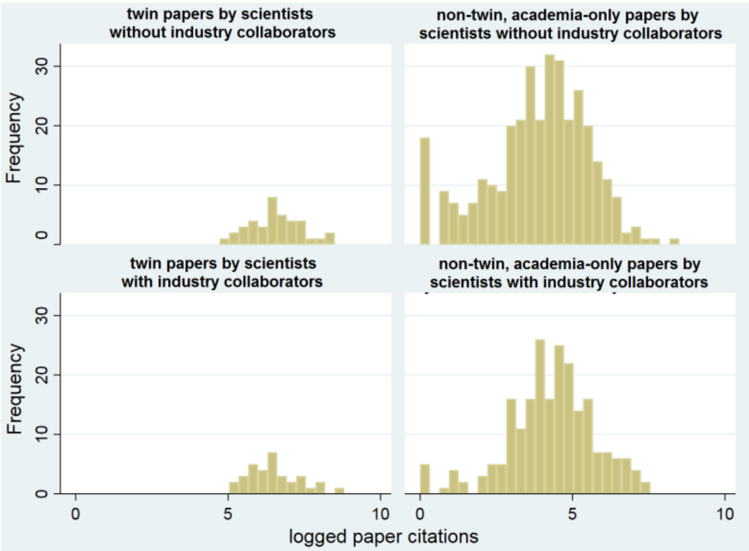

Figure 2. Patent Citation (0/1) Distribution of Twin and Non-Twin Papers Published by Corresponding Authors in the Same Year

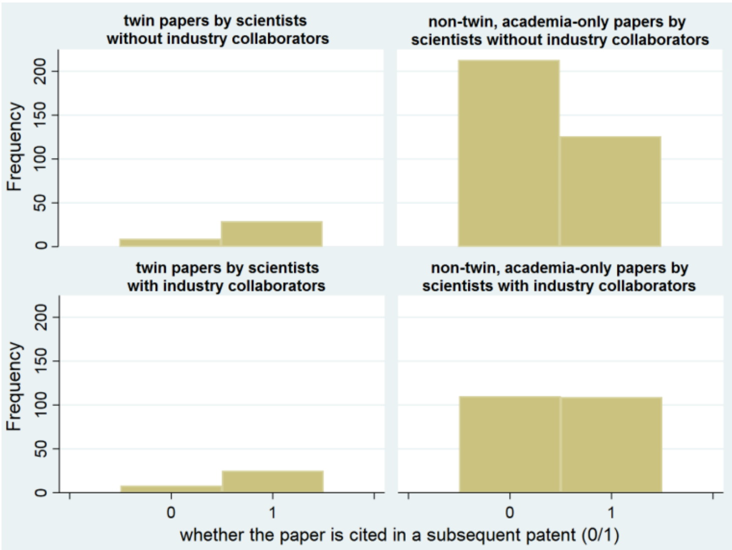

Figure 3. Overlap in Academic Citations between Academia-Only Twins and Non-Twins by Both Groups of Corresponding Authors

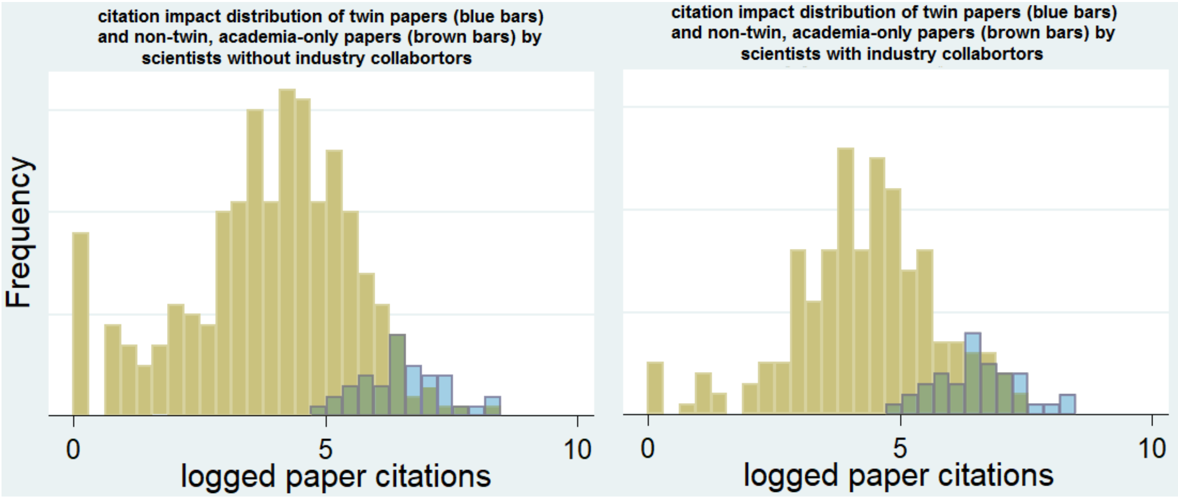

Table 1. Variable Definitions

Variable	Definition
Main author's follow-on papers	Number of papers by the main scientists of a simultaneous discovery citing the original paper that reports the discovery.
Main author's citation-weighted follow-on papers	Citation-weighted number of papers by the main scientists of a simultaneous discovery citing the original paper that reports the discovery.
Main author's follow-on patents	Number of patents by the main scientists of a simultaneous discovery citing the original paper that reports the discovery.
Main author's citation-weighted follow-on patents	Citation-weighted number of patents by the main scientists of a simultaneous discovery citing the original paper that reports the discovery.
Main author's total post-twin papers outside the simultaneous discovery line of research	Total number of papers by the main scientists of a simultaneous discovery that do not cite the original paper reporting the discovery.
Academia–industry collaboration	A dummy variable equal to 1 if the paper that reports a simultaneous discovery is a collaborative work between academia and industry; 0 otherwise.
Number of authors	The number of authors on the paper that reports the simultaneous discovery.
Main author's experience	Number of years since author's latest degree.
Main author affiliated with a U.S. institution	A dummy variable equal to 1 if the main author on the paper reporting a simultaneous discovery is affiliated with an institution in the U.S.; 0 otherwise.
Share of authors affiliated with U.S. institutions	Share of authors affiliated with U.S. institutions on the paper that reports a simultaneous discovery.
Share of female authors	Share of female authors on the paper that reports a simultaneous discovery.
Industry grant	A dummy variable equal to 1 if the paper reporting a simultaneous discovery is supported by an industry grant; 0 otherwise.
Discovery's commercial potential	Total number of patents that cite a simultaneous discovery.

Table 2. Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Max
Main author's follow-on papers	72	20.944	17.489	3	85
Main author's citation-weighted follow-on papers	72	3,041.583	4,414.692	66	26,560
Main author's follow-on patents	72	0.444	0.962	0	4
Main author's citation-weighted follow-on patents	72	0.583	1.392	0	7
Main author's total post-discovery non-twin papers	72	249.903	255.040	5	1365
Academia–industry collaboration	72	0.458	0.502	0	1
Number of authors	72	10.236	9.737	2	65
Main author's past paper count	72	113.153	115.926	1	447
Main author's past patent count	72	1.875	4.587	0	32
Main author's experience	72	16.239	8.496	-2	39
Main author affiliated with a U.S. institution	72	0.639	0.484	0	1
Share of authors affiliated with U.S. institutions	72	0.614	0.408	0	1
Share of female authors	72	0.298	0.185	0	0.667
Industry grant	72	0.264	0.444	0	1
Discovery's commercial potential	72	26.750	62.380	0	345

Table 3. Summary Statistics

Variable	Normalized past paper count	Normalized follow-on papers
Scientists with industry collaborators	0.826 (0.561)	0.940 (0.437)
Scientists without industry collaborators	0.967 (0.533)	0.868 (0.431)

Table 4. Impact of Academia–Industry Collaboration on Follow-On Publications and Patents

	Poisson Model			
	DV = Main author's follow-on papers	DV = Main author's citation-weighted follow-on papers	DV = Main author's follow-on patents	DV = Main author's citation-weighted follow-on patents
	(4-1)	(4-2)	(4-3)	(4-4)
Academia–industry collaboration (0/1)	0.313***	0.410***	-7.514*	-6.513**
	(0.118)	(0.126)	(4.051)	(2.954)
Number of authors	-0.018**	0.012	0.692***	0.655***
	(0.012)	(0.011)	(0.264)	(0.216)
Log(main author's past paper count+1)	0.158**	0.551***	-4.083	-3.068**
	(0.078)	(0.074)	(3.806)	(2.746)
Log(main author's past patent count+1)	0.034	-0.037	1.726	1.143
	(0.091)	(0.112)	(1.712)	(1.326)
Log(main author's academic experience+3)	0.323	-0.330	5.457	4.941
	(0.208)	(0.269)	(4.588)	(3.301)
Main author affiliated with a U.S. institution	0.040	1.280***	22.740	19.862
	(0.294)	(0.432)	(23.669)	(18.040)
Share of authors affiliated with U.S. institutions	-0.730**	-1.580***	-12.240	-10.522
	(0.333)	(0.422)	(19.194)	(15.206)
Share of female authors	-0.912	-1.499**	-16.803	-14.991
	(0.566)	(0.673)	(16.792)	(12.603)
Industry grant	-0.265	-0.208	7.723	6.732*
	(0.250)	(0.269)	(5.588)	(4.016)
Twin fixed effects	Yes	Yes	Yes	Yes
Observations	72	72	72	72
Log likelihood	-258.588	-14,593.686	-23.412	-24.043
Paper-twin FE	33	33	33	33

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Robustness Checks of Patent Results with Logit Regression

	Logit Model	Probit Model
	DV = Main author having a follow-on patent based on twin paper	DV = Main author having a follow-on patent based on twin paper
	(5-1)	(5-2)
Academia–industry collaboration (0/1)	-1.734	-1.086
	(1.240)	(0.724)
Controls	Yes	Yes
Twin fixed effects	Yes	Yes
Observations	16	16
Log likelihood	-9.172	-9.121
Paper-twin FE	7	7

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level.
*** p<0.01, ** p<0.05, * p<0.1

Table 6. Role of Commercial Potential and Industry Partner Type

	Poisson Model			
	DV = Main author's follow-on papers	DV = Main author's citation-weighted follow-on papers	DV = Main author's follow-on papers	DV = Main author's citation-weighted follow-on papers
	(6-1)	(6-2)	(6-3)	(6-4)
Academia–industry collaboration (0/1)	-0.198 (0.261)	-0.070 (0.279)	-0.033 (0.178)	-0.051 (0.213)
Academia–industry collaboration × Discovery has high commercial potential	0.631** (0.279)	0.529* (0.302)		
Academia–industry collaboration × Industry partner is established			0.549*** (0.214)	0.716*** (0.276)
Full set of controls	Yes	Yes	Yes	Yes
Twin fixed effects	Yes	Yes	Yes	Yes
Observations	72	72	72	72
Log likelihood	-248.811	-14,032.785	-249.623	-12,935.984
Paper-twin FE	33	33	33	33

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Robustness Tests with Negative Binomial and OLS Regressions

	Negative Binomial		OLS	
	DV = Main author's follow-on papers	DV = Main author's citation-weighted follow-on papers	DV = Main author's follow-on papers	DV = Main author's citation-weighted follow-on papers
	(7-1)	(7-2)	(7-3)	(7-4)
Academia–industry collaboration (0/1)	-0.508 (0.369)	-0.427 (0.297)	-0.517 (0.480)	-0.384 (0.343)
Academia–industry collaboration × Discovery has high commercial potential	0.870** (0.391)		0.934* (0.5343)	
Academia–industry collaboration × Industry partner is established		1.122*** (0.388)		1.133** (0.441)
Full set of controls	Yes	Yes	Yes	Yes
Twin fixed effects	Yes	Yes	Yes	Yes
Observations	72	72	72	72
Log likelihood	-581.368	-578.205		
R-squared			0.763	0.784
Paper-twin FE	33	33	33	33

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level. *** p<0.01, ** p<0.05, * p<0.1

Table 8. Likelihood of Collaborating with Industry Based on Main Author's Past Performance

	LOGIT MODEL	OLS MODEL
	DV = Academia–Industry collaboration	DV = Academia–Industry Collaboration
	(8-1)	(8-2)
Log(main author's past paper count+1)	-0.839**	-0.171
	(0.381)	(0.102)
Log(main author's past patent count+1)	0.358	0.069
	(0.456)	(0.147)
Log(main author's academic experience+3)	-0.187	-0.027
	(1.029)	(0.300)
Main author affiliated with a U.S. institution	0.621	0.134
	(0.870)	(0.252)
Industry grant	0.606	0.121
	(0.959)	(0.301)
Twin fixed effects	Yes	Yes
Observations	72	72
R-squared	0.110	0.162
Paper-twin FE	33	33

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level. *** p<0.01, ** p<0.05, * p<0.1

Table 9. Impact of Academia–Industry Collaboration on Main Author's Research Direction

	Poisson Model	
	DV = Main author's total post-twin papers outside the simultaneous discovery line of research	DV = Main author's total citation-weighted post-twin papers outside the simultaneous discovery line of research
	(9-1)	(9-2)
Academia–industry collaboration (0/1)	-0.042	0.094
	(0.096)	(0.088)
Full set of controls	Yes	Yes
Twin fixed effects	Yes	Yes
Observations	72	72
Log likelihood	-684.090	-46,197.167
Paper-twin FE	33	33

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level. *** p<0.01, ** p<0.05, * p<0.1

Table 10. Impact of Industry Sponsorship on Main Author's Research Output

	Poisson Model	
	DV = Main author's follow-on papers	DV = Main author's citation-weighted follow-on papers
	(10-1)	(10-2)
Industry grant (0/1)	-0.084	-0.030
	(0.084)	(0.109)
Full set of control	Yes	Yes
Twin fixed effects	Yes	Yes
Observations	1,151	1,151
Log likelihood	-4,607.276	-342,205.520
Paper-twin FE	540	540

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level. *** p<0.01, ** p<0.05, * p<0.1

Appendix

When Collaboration Bridges Institutions: The Impact of University–Industry Collaboration on Academic Productivity

Table A1- The Comparison between Non-Twin Papers of Scientists with Industry Collaborators and Those without Industry Collaborators

	POISSON MODEL	LOGIT MODEL	POISSON MODEL
	DV = NUMBER OF ACADEMIC CITATIONS RECEIVED	DV = WHETHER THE PAPER CITED IN A PATENT (=1, 0 OTHERWISE)	DV = NUMBER OF AUTHORS
	(A1-1)	(A1-2)	(A1-3)
Academia–industry collaboration on the twin published in the same year (0/1)	0.031 (0.154)	0.153 (0.197)	0.102 (0.086)
Twin fixed effects	Yes	Yes	Yes
Observations	447	372	447
Log-likelihood	-45572.015	-217.310	-1515.973

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level. *** p<0.01, ** p<0.05, * p<0.1

Table A2- The Comparison of the Composition of Authorship Team within Twins

	POISSON MODEL			
	DV = NUMBER OF FACULTY	DV = NUMBER OF RESEARCH STAFF (GRAD STUDENTS & POSTDOCS)	DV = NUMBER OF AUTHORS W/ UNIDENTIFIED STATUS	DV = NUMBER OF TOTAL PRE-DISCOVERY PUBLICATIONS OF THE AUTHORSHIP TEAM
	(A2-1)	(A2-2)	(A2-3)	(A2-4)
Academia–industry collaboration (0/1)	0.249 (0.170)	0.055 (0.136)	0.096 (0.127)	0.090 (0.201)
Twin fixed effects	Yes	Yes	Yes	Yes
Observations	72	72	72	72
Log-likelihood	-103.990	-82.281	-147.784	-5,299.343

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level.
*** p<0.01, ** p<0.05, * p<0.1

Table A3- Robustness Checks with the Limited Control Set

	POISSON MODEL					
	DV = MAIN AUTHOR'S FOLLOW-ON PAPERS	DV = MAIN AUTHOR'S CITATION-WEIGHTED FOLLOW-ON PAPERS	DV = MAIN AUTHOR'S FOLLOW-ON PAPERS	DV = MAIN AUTHOR'S CITATION-WEIGHTED FOLLOW-ON PAPERS	DV = MAIN AUTHOR'S FOLLOW-ON PAPERS	DV = MAIN AUTHOR'S CITATION-WEIGHTED FOLLOW-ON PAPERS
	(A3-1)	(A3-2)	(A3-3)	(A3-4)	(A3-5)	(A3-6)
Academia-industry collaboration (0/1)	0.226* (0.118)	0.267** (0.133)	-0.357 (0.277)	-0.243 (0.305)	-0.081 (0.161)	-0.086 (0.198)
Academia-industry collaboration × Discovery has high commercial potential			0.736** (0.300)	0.584* (0.335)		
Academia-industry collaboration × Industry partner is established					0.485** (0.221)	0.502** (0.249)
Log(main author's past paper count+1)	0.233*** (0.083)	0.519*** (0.083)	0.205*** (0.069)	0.506*** (0.106)	0.238*** (0.052)	0.494*** (0.097)
Main author affiliated with a U.S. institution	-0.409** (0.083)	-0.330* (0.083)	-0.402** (0.161)	-0.332 (0.204)	-0.443** (0.175)	-0.385** (0.189)
Twin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	72	72	72	72	72	72
Log likelihood	-282.218	-20,592.058	-267.057	-19,753.719	-273.451	-19,553.834
Paper-twin FE	33	33	33	33	33	33

Robust standard errors are reported in parentheses. The level of analysis is at the twin-paper level.
*** p<0.01, ** p<0.05, * p<0.1