



City Research Online

City St George's, University of London

Citation: Banal-Estanol, A., Jofre-Bonet, M. & Lawson, C. (2015). The Double-Edged Sword of Industry Collaboration: Evidence from Engineering Academics in the UK. *Research Policy*, 44(6), pp. 1160-1175. doi: 10.1016/j.respol.2015.02.006

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/13878/>

Link to published version: <https://doi.org/10.1016/j.respol.2015.02.006>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

The Double-Edged Sword of Industry Collaboration: Evidence from Engineering Academics in the UK

Albert Banal-Estañol,^{ab*} Mireia Jofre-Bonet^a and Cornelia Lawson^{cd}

^a City University London, Department of Economics, Northampton Square, London, EC1V 0HB, UK

^b Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain

^c BRICK, Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri (Turin), Italy

^d University of Nottingham, School of Sociology and Social Policy, University Park, Nottingham, NG7 2RD, UK

* Corresponding author: Albert Banal-Estañol, Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain, email: albert.banalestanol@upf.edu, Tel.: +34 93 542 2871, Fax: +34 93 542 1746

Email addresses: albert.banalestanol@upf.edu (Banal-Estañol), mireia.jofre-bonet@city.ac.uk (Jofre-Bonet), cornelia.lawson@nottingham.ac.uk (Lawson).

Abstract

This paper studies the impact of university-industry collaboration on academic research output. We analyze the channels through which the *degree* of industry collaboration may be affecting research output. We exploit a unique longitudinal dataset on all the researchers in all the engineering departments of 40 major universities in the UK for the last 20 years. We use an innovative measure of collaboration based on the fraction of public research grants that include industry partners. Our empirical findings corroborate that the relationship between collaboration degree and publication rates is curvilinear, and shed some light on the selection mechanisms at work. Our results are robust to several econometric methods, measures of research output, and subsamples of academics.

Keywords: Industry-science links, research collaboration, basic vs. applied research.

JEL codes: O3, L31, I23

1 Introduction

In a modern economy transforming scientific research into competitive advantages is essential. In the US, extensive collaboration between universities and industry, and the ensuing transfer of scientific knowledge, is viewed as one of the main contributors to the successful technological innovation and economic growth of the past three decades (Hall, 2004). At the same time, according to a European Commission report (1995), insufficient interaction between universities and firms in the EU has been one of the main factors for the EU's poor commercial and technological performance in high-tech sectors. Nowadays, increasing university-industry collaboration is a primary policy aim in most developed economies.¹

The increased incentives (or, as some say, *pressure*) to collaborate with industry may have controversial side effects on the production of scientific research itself. Nelson (2004), among many others, argues that the existence of industry involvement might delay or suppress scientific publication and the dissemination of preliminary results, endangering the "intellectual commons" and the practice of "open science" (Dasgupta and David, 1994). Florida and Cohen (1999) argue that industry collaboration might come at the expense of basic research: growing ties with industry might be affecting the choice of research projects, "skewing" academic research from a basic toward an applied approach.

Academics that contribute to knowledge and technology transfer, on the other hand, maintain that the existence of industry collaboration complements their own academic research by securing funds for graduate students and lab equipment, and by providing them with ideas for their own research (Lee, 2000, Agrawal and Henderson, 2002). Siegel et al. (2003), for example, report that "[s]ome scientists explicitly mentioned that these interactions improved the quantity and quality of their basic research." Ideas sourced from industry may thus expand traditional research agendas (Rosenberg, 1998), benefitting the overall scientific performance of researchers.

These opposing claims raise a long-standing question for academic research: Does collaboration with industry increase or decrease publication rates? Previous research on this issue has mostly used patenting as a measure of industry collaboration (see Geuna and Nesta, 2006, and Baldini, 2008, for reviews). The evidence is somewhat mixed, ranging from the negative effects of patenting on research output reported in surveys of

¹ In the 1980s, the US introduced a series of structural changes in the intellectual property regime accompanied by several incentive programs, designed specifically to promote collaboration between universities and industry (Lee, 2000; Mowery et al., 2001). Almost 30 years on, many elements of the US system of knowledge transfer have been emulated in many other parts of the world (see e.g., the UK Government's White Paper "The Future of Higher Education," 2003).

academic scientists (Blumenthal et al., 1996), to no effect in some of the econometric studies (Agrawal and Henderson, 2002) to even a positive relationship in some of the recent evidence (Azoulay et al., 2009; Breschi et al., 2008; Fabrizio and DiMinin, 2008; Stephan et al., 2007; van Looy et al., 2006).

This paper argues that academic research output is affected not only by the *existence* of links with the industry but also by the *degree of industry collaboration*, i.e., by the proportion of (or the share of time spent on) projects with industry involvement.² By exploring the channels through which industry collaboration affects publications, our conceptual framework suggests that the relationship between collaboration degree and publication rates is neither increasing nor decreasing but is curvilinear, described by an inverted U-shaped curve. As a result, research output shall be maximized at *intermediate* degrees of collaboration, i.e., when the industry is involved in some but not in all the projects of the academic.

Our empirical analysis uses an innovative measure of (degree of) industry collaboration based on the fraction of publicly funded research projects which include industry partners.³ In contrast to patents, this measure is continuous in nature, and so is able to proxy not only for the existence but also for the degree of industry collaboration. In addition, collaborative links through joint research, consulting or training arrangements are more widespread (D’Este and Patel, 2007) and are more important knowledge transfer channels than patents, licenses, and spin-offs, according to both academics (Agrawal and Henderson, 2002) and firms (Cohen et al., 2002). Data on research collaborations also provide a more continued assessment of the level of interaction with industry than measurements based on the number of patents. Possibly due to the lack of comparable data, the literature has paid little attention to these more collaborative forms of university-industry interaction.

Our measure of collaboration is constructed exploiting comprehensive information from the main UK government agency for funding in engineering, the Engineering and Physical Sciences Research Council (EPSRC), which distinguishes between collaborative and non-collaborative research grants based on the involvement of industry partners. In addition to research funds, we compiled a unique, longitudinal dataset containing

² Our notion of “degree of (industry) collaboration” is inspired by the notion of “degree of (research) collaboration” used in bibliometric studies. As shown by Subramanyam (1983), the degree of research collaboration is usually defined as the number of multicoauthored papers out of the total number of papers (single and multicoauthored).

³ The presence of industry partners in public research grants might not be a perfect proxy for the degree of collaboration with industry, as there are other channels of interaction. The inclusion of private firms as partners in these grants, however, is highly correlated with obtaining direct funding from industry (Meissner, 2011).

research output (publications), patents, and other individual characteristics for all academics employed in all the engineering departments of 40 major UK universities between 1986 and 2007. Since our dataset contains the majority of academic engineers in the UK, our results are not driven by the most successful or academic inventors alone. In fact, we can test whether the effects differ across observed categories of researchers.

Still, the observed degrees of collaboration are not exogenously determined, but are the result of individual and *mutual* choices in a two-sided market of academics and firms (Mindruta, 2013; Banal-Estañol et al, 2014). Unobserved characteristics of the researchers may affect not only their degree of collaboration but also their academic productivity, thereby influencing the shape of the collaboration-publication relationship. Our conceptual framework analyzes the potential selection mechanisms at work, and the direction of the biases one might incur if these mechanisms were ignored in the estimation. As empirical strategy, to deal with the endogeneity problem, we use fixed effects and instrumental variable techniques and, to take into account the dynamic nature of the publications (e.g., Arora et al., 1998; Agrawal and Henderson, 2002), we use a dynamic panel data approach. By comparing instrumented and non-instrumented regressions we shall also shed more light on the selection mechanisms in place.

The paper is organized as follows. In section 2 we provide the conceptual framework. In section 3 we describe the dataset and in section 4 we introduce our empirical strategy. Section 5 presents our results. Section 6 discusses and concludes.

2 Conceptual framework

Our conceptual framework is based on the idea that the degree of industry collaboration of an academic affects the main determinants of her scientific output, namely, (i) the quality and quantity of her ideas; (ii) the time and attention she can devote to developing these ideas and transforming them into papers; (iii) the amount of resources she has available;⁴ and (iv) the existence of constraints on the scope and/or in the dissemination of research results (Stephan, 1996, 2012). In the first subsection below, we discuss the channels through which the degree of industry collaboration may be affecting research output. In the second subsection, we consider the characteristics of the academics that may be affecting both their observed degrees of collaboration and their

⁴ The availability of research funding is important for scientists in all academic disciplines, but especially in resource-intensive fields such as engineering (Stephan, 1996, 2012). Several recent studies have documented a positive impact of public grants on research performance (Jacob and Lefgren, 2011; Benavente et al., 2012).

research output. We describe the potential selection mechanisms at work, allowing us to identify potential biases in the estimation.

2.1 Effects of the degree of industry collaboration on research output

Collaboration with industry can boost research output for at least two reasons. First, collaboration can expand academics' research agendas and improve the pool of research ideas (Rosenberg, 1998). Mansfield (1995) shows that a substantial number of publicly sponsored research projects stem from industrial problems encountered in consulting. Collaboration helps academics gain new insights for their own research and test the practical application of their theoretical ideas (Lee, 2000). The generation and/or refinement of ideas through puzzle-solving may in turn improve research outcomes because the resulting ideas can be transformed into more and/or better academic papers.

Second, industry collaboration can expand the availability of financial resources. According to survey evidence in Lee (2000), two of the most important reasons for academics to collaborate are to secure funds for graduate students and lab equipment, and to supplement funds for their academic research. In recent years, industry has been identified as an even more important source of funding for academic research. Private financial support is important in light of progressive declines in direct government funding (OECD, 2010) and of more competitive research environments (Stephan, 2012).

Nevertheless, an increase in the degree of industry collaboration is not necessarily positive for research output for at least five reasons. First, although collaboration may enhance the pool of research ideas, there might be decreasing returns to scale associated to the generation of these ideas. Hottenrott and Lawson (2014) show that research units that receive larger shares of funding originating from industry are also more likely to develop ideas stemming from private partners, suggesting that the pool of ideas for higher degrees of collaboration is indeed larger. But, at high degrees of collaboration, ideas at the margin may not be of the same quality, or may not even be worth pursuing, and may thus not result in the same increase in publication rates as the ideas obtained at low degrees of collaboration.

Second, a high number of ideas, conceivably available from higher degrees of collaboration, may also create attention problems. According to attention-based theories of the firm, decision-makers in any organization need to "concentrate their energy, effort and mindfulness on a limited number of issues" (Ocasio, 1997). As argued by Laursen and Salter (2006) in the context of innovative firms, if there are too many ideas, few of

them receive the required level of time or effort to be developed seriously. As a result, high levels of collaboration may generate many ideas but few academic papers.

Third, collaborative projects usually require time for coordination, organization, and interaction. Collaboration with a private partner may also come with “strings attached” in the form of academic consulting or commercial activities. The general duties of the academics, and research in particular, might be compromised by an increase in the time allocated to development, consulting or commercialization (Florida and Cohen, 1999), thus reducing scientific publication.

Fourth, collaboration may affect the selection of topics and methodologies (Florida and Cohen, 1999). As argued in Trajtenberg et al. (1997), industry research and development tends to be directed at commercial success, while university research generally focuses on solving fundamental scientific questions. Thus, research that appeals to industry partners may not necessarily be close to the research frontier (Rosenberg and Nelson, 1994), and may be less likely to result in (top) academic publications. This is especially the case for academics with high degrees of collaboration who may get locked in service provision for industry (Meyer-Krahmer and Schmoch, 1998).

Finally, firms’ commercial interests might impose constraints on the publication activity of collaborating academics, especially those that collaborate extensively. Firms’ commercial interests may push firms to include non-disclosure clauses that delay or suppress scientific publication (Nelson, 2004). Czarnitzki et al. (2014) indeed find empirical evidence that the percentage of researchers that complain about secrecy and publication delay is larger for researchers sponsored by industry.

As summarized in the upper part of Figure 1, collaboration can have positive and negative effects on the factors driving academic research. The relative magnitudes of these effects and their ultimate impact on publications change with the degree of collaboration. From our discussion, we expect the negative effects to be relatively more important and, thus, to dominate for high degrees of collaboration while the positive ones shall dominate for low degrees of collaboration. We therefore anticipate the relationship between collaboration degree and publication rates to have an inverted U-shape and research output to be maximized at intermediate degrees of collaboration, i.e., when industry is involved in some but not all the projects of the academic.

2.2 Mechanisms influencing the collaboration-research output relationship

This subsection discusses other mechanisms that may affect the relationship between industry collaboration and research output described in the previous subsection. We argue that the degrees of collaboration are not exogenously determined, but are the result of individual and *mutual* choices in a two-sided ‘market’ of academics and firms (Mindruta, 2013; Banal-Estañol et al, 2014).⁵ Observed and unobserved characteristics of the researchers (such as seniority, ability or skills) may affect not only their degree of collaboration but also their academic productivity, thereby influencing the shape of the collaboration-publication relationship. In this subsection, we focus on the unobserved characteristics and describe the potential mechanisms at work, by ‘type’ of researcher, and the direction of the biases one might incur if these mechanisms were ignored when estimating the causal effect of degree of collaboration on research output.

Notice first that there are inherent characteristics of the researchers that make them both more likely to publish and more likely to find a good partner. For example, *ability* or talent is important for research but it is also highly valued by firms when searching for research partners (see e.g., Blumenthal et al., 1996). As a result, highly able academics may end up having higher degrees of collaboration and better partners than less-able academics. As argued by Blumenthal et al. (1986), “the most obvious explanation for [the] observed positive relation [between collaboration and publication] is that companies selectively support talented and energetic faculty who were already highly productive.” Through their involvement with more attractive partners, they may end up producing even more academic papers. As shown in Banal-Estañol et al. (2013), collaborative projects generate more publications than non-collaborative ones if and only if the industrial partners are highly productive. As a result, if one does not take into account the presence of researcher “fixed effects” such as ability, the estimation may generate positive biases on the effect of high degrees of collaboration on publications.

Time-invariant individual characteristics may also interact with time-variant ones and generate other biases. First, some of the *talented academics* may develop *stronger preferences* for industry collaboration and *better networking skills*, thereby enjoying higher-quality interactions and more knowledge of the private sector. We expect these

⁵ Firms also weigh the benefits and costs of collaborating with academic partners (Henderson et al., 1998; Salter and Martin, 2001; Cohen et al., 2002; Link and Scott, 2005; Laursen et al., 2011). Firms report to collaborate to get access to new university research and discoveries (Lee, 2000), but are also concerned with the organizational and institutional structure, and the existence of the open science culture, in academia (Dasgupta and David, 1994). An individual firm’s decision to collaborate depends on its absorptive capacity (Veugelers and Cassiman, 2005), its size, and whether it adopts an open search strategy (Mohnen and Hoareau, 2003; Laursen and Salter, 2004).

“industry-savvy” researchers to have higher degrees of collaboration and, at the same time, to choose and to be chosen by better industrial partners because they screen better and they are more appealing academic partners. Thanks to their involvement with more attractive industry partners, these academics are again predicted to produce more academic papers than those at low degrees of collaboration. In sum, more engaged researchers may have both higher degrees of collaboration and – because they end up being matched with better partners – more research output than researchers lacking collaboration taste and networking skills. We may thus observe a positive bias at the high-end of the degree of collaboration (those more likely populated by industry-savvy researchers) and, symmetrically, a negative bias at low degrees of collaboration (populated by non-industry-savvy researchers).

Second, some of the *talented academics* may become *more output-driven* in their collaboration choices. We expect these researchers to publish more and, at the same time, prefer intermediate degrees of collaboration, as they are more likely to identify and ponder the trade-offs described in the previous subsection (e.g., new ideas vs. time constraints). That is, these highly able, output-driven researchers may end up having mid-range degrees of collaboration, while less able, less output-driven academics, may end up having insufficient or excessive degrees of collaboration that negatively affect their academic performance. In addition, the highly able will also end up partnering with more attractive partners, producing even more papers. We may thus observe a positive bias on the mid-range degrees of collaboration and a negative bias in both the low- and high-ends of the degree of collaboration.

Third, some of the *talented researchers* may also become more *selective* over time, given the increased pool of potential industry projects they have available. While collaboration with highly able scientists is the most beneficial for firms, companies find academic partners across a whole quality-range of researchers and departments, with the majority of industry funding going to universities of medium research quality (Mansfield, 1995). Goldfarb (2008) argues that star-scientists have many more opportunities for funding but it is the less-able researchers who typically engage in programs sponsored by mission-oriented agents such as firms. Less-able academics may have to accept any industry support in order to maintain or increase their funding (Carayol, 2003). In sum, talented, successful researchers may be in a position to be more selective and engage only in collaborations with high-quality industry partners. Due to the existence of these types of academics, we may observe a positive bias at the low-end

of the degree of collaboration and a corresponding negative bias in the mid- and high-ends.

The selection mechanisms described in this section are summarized in the lower part of Figure 1. As explained in more detail in the empirical strategy section, these mechanisms have fundamental implications for any attempt to estimate the relationship between the degree of industry collaboration and research output. The degree of collaboration is “endogenous,” and it is affected by observed and unobserved individual characteristics that also affect research output. Our empirical approach shall control for time-invariant and time-variant *observed* characteristics. We shall also take into account the existence of unobserved characteristics influencing the selection mechanisms at work. Not doing so would result in a biased estimate of the impact of degree of collaboration on research output. In our empirical strategy section, we explain how we address these endogeneity concerns.

3 Data

In this section, we provide a detailed account of how we created our dataset. For this study, we built a unique longitudinal dataset containing individual characteristics, publications, research funds, and patents for all researchers employed in all the engineering departments of 40 major UK universities between 1986 and 2007 (see Table 1 for a list of the universities). Through the British Library, we searched for university calendars and prospectuses providing detailed staff information for all the universities with engineering departments in the UK.⁶ Our final sample contains all the universities that had calendar information available, including all the universities that are members of the prestigious Russell Group, a coalition of 24 research-intensive UK universities, as well as 16 other comprehensive or technical universities.⁷

We retrieved the academics’ names and ranks for all the years from 1986 to 2007. We focused on academic staff carrying out both teaching and research and did not consider research officers or teaching assistants. We followed the researchers’ career paths

⁶ By Act of Parliament, the British Library is entitled to receive a free copy of every item published in the UK. These data were supplemented with information from the Internet Archive, a not-for-profit organization that maintains a free Internet library committed to offering access to digital collections. Their collection dates back to 1996 and enabled us to retrieve information from outdated Internet sites.

⁷ We identified the initial set of engineering departments from the 1996 and 2001 Research Assessment Exercises (RAEs). We did not find staff information for eight institutions which are similar to the 16 non-Russell group universities in our sample. We did not consider any of the 39 post-92 universities either, as these were not full research institutions for all the years considered in our analysis. We also excluded the Open University and Cranfield University which, as distance and postgraduate institutions, respectively, have a very different structure.

between the different universities by matching names and subject areas and by checking the websites of the researchers. Academics leave (and join or rejoin) our dataset at different stages in their career, when they move to (or from) abroad, industry, departments other than engineering (e.g., chemistry, physics), or universities that are not part of our dataset, resulting in an unbalanced panel. They represent the basis for our data collection and enable us to retrieve information on publications, research funds, and patents.

Our final sample contains information on 3,991 individuals. The final sample excludes all inactive researchers (those with neither publications nor funds during the entire sample period) and researchers who were present for less than six consecutive years so that all of our (stock) variables could be created.⁸ We describe below our sources and measures of research output (our dependent variable), degree of industry collaboration (our main independent variable), as well as funding, patents, and other individual characterizing variables. We provide summary statistics in the first panel of Table 2.

Research output. Data on publications were obtained from the ISI Science Citation Index (SCI). The number of publications in peer-reviewed journals is not the only measure but is the best recorded and the most accepted measure for research output as they are essential for gaining scientific reputation and for career advancements (Dasgupta and David, 1994). We collected information on all the articles published by researchers in our database while they were employed at one of the institutions in our sample. Most entries in the SCI database include detailed address data that allowed us to identify institutional affiliations and unequivocally assign articles to individual researchers.⁹

As a main measure of research output for each researcher in each year, we use the normal count of publications ($count_{it}$), i.e., the number of publications in t on which researcher i is named as an author. Publication counts, however, might be misleading for articles with a large number of authors and may not reflect a researcher's effective productivity. Therefore, we also use the co-author-weighted count of publications (*co-author weighted count* $_{it}$), which we obtain by weighting publications by the inverse of the publication's number of co-authors.

⁸ Estimations considering a shorter time window of just three consecutive years are used in the robustness checks. The descriptive statistics as well as the empirical results are very similar to those of the main estimation.

⁹ Publications without address data had to be ignored. However, we expect this missing information to be random and to not affect the data systematically.

We also separate the count of publications by type of academic research (basic or applied). To construct these measures, we use the Patent Board (formerly CHI) classification (version 2005), developed by Narin et al. (1976) and updated by Kimberley Hamilton for the National Science Foundation (NSF). Based on cross-citation matrices, it characterizes the general research orientation of journals, distinguishing between (1) applied technology, (2) engineering and technological science, (3) applied and targeted basic research, and (4) basic scientific research. Godin (1996) and van Looy et al. (2006) reinterpreted the categories as (1) applied technology, (2) basic technology, (3) applied science, and (4) basic science; and grouped the first two as “technology” and the last two as “science.” We use the normal count of publications in each of these two categories and denote them “applied” (*applied count_{it}*) and “basic” (*basic count_{it}*), respectively. Due to the applied character of the engineering sciences, 74% of all publications are applied.

Degree of industry collaboration. Our measure of industry collaboration is based on grants awarded by the Engineering and Physical Sciences Research Council (EPSRC), the main UK government agency for research in engineering and the physical sciences, and by far the largest provider of funding for research in engineering (more than 50% of overall third-party funding). The EPSRC encourages (but does not require) academic researchers to find private partners for their research projects. As defined by the EPSRC, “Collaborative Research Grants are grants led by academic researchers, but involve other partners.” Partners generally contribute either cash or “in-kind” services to the full economic cost of the project.¹⁰

We obtained information on all the grants awarded since 1986. For each grant, we collected the start year, duration, total amount of funding, names of principal investigator (PI) and co-investigators, grant-receiving institution, and names of partner organizations, if any. In order to construct our proxy for the *degree of collaboration with industry* we use the presence of private partners. Our variable, which we name *fraction of EPSRC funding with industry_{it}*, represents the fraction of collaborative EPSRC funds of an individual *i* in the five previous years (i.e., between *t-4* and *t*). We use a five-year window to reflect the profile of an academic in terms of her past stream of funding.

To be precise, this variable was constructed as follows. We divided the total monetary income of each grant between the PI and her co-investigator(s). We took into account the

¹⁰ The EPSRC does not favor specific types of academic research output. Both collaborative and non-collaborative grants are awarded based on peer-review and monitored through end-of-award reports. There are, however, no specific measures for evaluating the success of the knowledge exchanges between science and industry.

participation of all investigators but positively discriminated PIs by assigning them half the grant value and splitting the remaining 50% among the co-investigators. PIs were assigned a major part as they are expected to lead the project. We additionally spread the grant value over the award period.¹¹ This was done in order to account for the ongoing benefits and costs of the project and to mitigate the effect of focusing all the funds at the start of the project. Finally, for each year and for each researcher we computed the fraction of “collaborative” funds in the last five years, i.e., those that included one or more industry partners, over all EPSRC funds received in the same period.

Funding. As discussed in the *conceptual framework* section, the availability of funds is an important factor for research output. We therefore created an indicator variable (*had some EPSRC funding_{it}*) that takes value 1 if academic *i* received any EPSRC funding (collaborative or not) in the five previous years and 0 otherwise. We used the indicator instead of the funding level because the latter is discipline-specific (some disciplines require more funding than others). As for industry collaboration, we used a five-year window to reflect the academic’s profile.

Patents. In the *conceptual framework*, we argued that the commercialization of research results might impose constraints on publication activity. Our analysis should therefore control for patent activity. By including patents, we also separate the effect of patenting from the effect of collaborating with industry, as defined above. Prior research has considered patenting itself to be an indicator of a researcher’s involvement with industry. As a result, the benefits and costs of collaboration might also appear through the patent channel.

We obtained patent data from the European Patent Office (EPO) database. We collected those patents that identify the aforementioned researchers as inventors and were filed while they were employed at one of the 40 institutions. We not only consider patents filed by the universities themselves, but also those assigned to third parties, e.g., industry or government agents (as shown by Lawson, 2013b, 52% of academic patents in the UK are not owned by the university). The filing date of a patent was recorded as representing the closest date to invention. Since the filing process can take several years,

¹¹ If the grant lasted two years, we split it equally across those two years. If it lasted three or more years, the first and the last years (which are assumed to not represent full calendar years) received half the share of an intermediate year.

we were only able to include patents published by 2007, hence filed before 2005.¹² The EPO only covers a subsample of patents filed with the UK Intellectual Property Office (UKIPO). Nevertheless, the patents that are taken to the EPO are those with a higher economic potential and/or quality (Maurseth and Verspagen, 2002) and have been used in the past to analyze academic patenting in Europe (see Lissoni et al., 2008).

To measure the impact of patenting on the timing of publications, we use a dummy variable indicating whether the academic i filed any patent in the same year ($patent_{it}$), or in the two years preceding the publication ($patent_{it-1}$ and $patent_{it-2}$). Researchers in Europe, unlike the US, cannot benefit from a “grace period” and hence have to withhold any publication related to the patent until the patent application is filed. We therefore expect a lag of up to two years between invention and publication in a journal.

Individual characteristics. Research output might be linked to the researcher’s personal attributes such as sex, age, education, and academic rank. Academic Rank_{it} is the only time-variant observable characteristic in our dataset. Thus, we incorporate information on the evolution of researchers’ academic status from lecturer to senior lecturer, reader, and professor. Lecturer corresponds to an assistant professor in the US, whereas senior lecturer and reader would be equivalent to an associate professor.

We also include, as an additional time-variant characteristic of an academic at a given point in time, her past publications. Indeed, as argued by Stephan (1996), there is a “cumulative advantage” in science that results in a dynamic relationship between past and present publication output.

Interaction variables. The effect of degree of industry collaboration on research output might differ across observed categories of academics. In order to investigate whether the relationship differs between academics who have certain observed characteristics, we create indicator variables that reflect (i) being above or below the median lifetime share of publications with industry co-authors; (ii) being above or below the median lifetime share of collaborative grants; (iii) being above or below the median amount of funding during the previous five years; (iv) belonging to the selected Russell group of universities; and, (v) being at an earlier stage of their careers as opposed to being senior researchers (professors). We allow researchers to change groups when they

¹² Just like previous studies (see e.g., Fabrizio and DiMinin, 2008), data construction requires a manual search in the inventor database to identify entries that were truly the same inventor and exclude others with similar or identical names. This was done by comparing the address, title, and technology class for all patents potentially attributable to each inventor. The EPO database is problematic in that many inventions have multiple entries. It was thus necessary to compare priority numbers to ensure that each invention is only included once in our data.

change universities or are promoted. Panels 2 to 6 in Table 2 present the descriptive statistics of the two main variables of interest for these subsamples.

4 Empirical strategy

This section describes the econometric specification of our model, the methods we use to estimate it, and the instrumental variables we exploit in order to do so.

4.1. Econometric specification

According to our conceptual framework, the relationship between the degree of collaboration and the publications of an academic can be curvilinear. Indeed, increasing the degree of collaboration with industry can boost research output, as collaboration may improve the pool of research ideas and expand the availability of financial resources. But, at high degrees of collaboration, it may also be negative, because there might be decreasing returns associated to the generation of ideas, extensive industry involvement may impose time constraints and cause attention problems, and commercial objectives and strategic behavior may push industrial partners to impose constraints on the selection of topics and methodologies and the dissemination of research results.

So we estimate a model where academic research output is a quadratic function of the degree of collaboration with industry. Additionally, we make use of the variables described in the previous section to control for other factors, other than the degree of collaboration, which may also affect research output. We include, for example, the ability of having raised EPSRC research funds, which proxies for other resources the researcher may have available. We also incorporate patent indicator variables to account for the existence of other constraints on the scope and/or in the dissemination of research results. We also include past academic output to account for other factors affecting the pool of ideas of the researcher and her ability to transform them into papers. To proxy for seniority and the existence of other time constraints, we also include her academic rank.

Accordingly, we formulate the following empirical model:

$$y_{it} = \sum_{j=1}^2 \alpha_j y_{i(t-j)} + \beta_1 hf_{it-1} + \beta_2 fin_{it-1} + \beta_3 fin_{it-1}^2 + \sum_{k=0}^2 \gamma_k p_{i(t-k)} + \delta x_{it-1} + \mu_i + v_{it},$$

where y_{it} stands for academic i 's research output at time t (either *count_{it}*, *co-author weighted count_{it}*, *applied count_{it}* or *basic count_{it}*); hf_{it-1} is the indicator variable *had some EPSRC funding_{it-1}*; fin_{it-1} is the *fraction of EPSRC funding with industry_{it-1}*; p_{it} is the indicator variable *patent_{it}* for having filed at least one *patent* at time t ; and, x_{it-1} is a vector

of other time-variant individual characteristics including *Academic Rank*_{*it*-1} and year. All the independent variables except the patents are lagged because of the publication lead time. The error term contains two sources of error: the academic *i*'s fixed effect term μ_i and a disturbance term v_{it} . Since the distributions are highly skewed, we take logarithms of both the research output and degree of industry collaboration variables. As these figures contain zero values, we add the unit before we take logarithms.

4.2. Empirical methods

The presence of time-invariant individual factors, μ_i , in the error term produces correlation among the individual errors across different periods of time, making Ordinary Least Squares (OLS) inefficient and yielding incorrect standard errors. In addition, as explained in the conceptual framework, there are inherent characteristics of the researchers (e.g., ability) that make them both more likely to publish and more likely to find good partners and, therefore, to have higher degrees of collaboration. This creates problems of endogeneity for our main variables of interest. Thus, we first estimate our model using a Generalized Least Squares with a fixed effects estimator (GLS FE).

But still, correcting for fixed effects alone may not be sufficient to address the biases introduced by the selection mechanisms described in the conceptual framework. There are time-variant unobserved individual characteristics, such as becoming more industry-savvy, more output-driven, and/or more selective. These time-variant individual unobserved traits may affect both research output and degree of collaboration causing the latter independent variable to be again endogenous. Fixed Effects based methods are not sufficient in dealing with this source of time-variant endogeneity. Thus, to obtain consistent estimates of the coefficients of interest, we use an estimator that *instruments* our collaboration measures with *instrumental* variables that affect industry collaboration but not research output directly (GLS FE IV). Below, we discuss which instruments we use.

GLS models, though, cannot correct for the error autocorrelation created by the inclusion of past research outputs (i.e., lagged dependent variables) as explanatory variables. Therefore, we also estimate a dynamic Generalized Method of Moments (GMM) panel data model: the Arellano-Bond GMM estimator (GMM AB) (Arellano and Bond, 1991; Blundell and Bond, 1998). This estimator transforms the model into first differences and eliminates the individual effects – and, thus, the cause of the autocorrelation across time periods. Lagged dependent variables are instrumented with not only the exogenous variables described in the next subsection but also with deeper

lags of the independent and dependent variables which are remote enough in the past so that their correlation with current publications has been dissipated. Indeed, we make sure that the GMM models satisfy both the Autocorrelation test and the Sargan test of over-identifying restrictions.¹³ In our case, the required depth of the lags is three periods, whereas the funding history goes back five years. To make sure that this is not the cause of further hidden autocorrelation, and as a robustness check, we also estimate our model using a funding stock variable based on only *two years* of funding, as opposed to five.

4.3. Instrumental variables

We instrument the *fraction of EPSRC funding with industry* variable using the *economic activity of the area* and the *overall share of industry funding of the department*. Economic activity of the area is approximated by the yearly number of manufacturing firms, as listed in the COMPUSTAT database, in the own and adjacent postcodes of the university where the academic works. The share of funding from industry received by the whole department is obtained from Research Assessment Exercise (RAE) data, which provide information on the amount of research funds received by each department in the UK, decomposed by source (public, private, and other funding) for the years 1993 to 2007. We also instrument the variable *had some EPSRC funding*, using the *aggregate amount of funding received by the department*, based on the same RAE data.

Our instruments for the degree of collaboration assume that local economic activity and the overall industry involvement of the department do not affect individual research output but do have an impact on the individual's opportunity to collaborate with firms. Similarly, total funding of the department, our instrument for *had some EPSRC funding*, should not affect individual research output but should have an impact on the individual's ability and opportunity to obtain funds. Our instruments are jointly significant in the first stage regressions. The residuals-based Smith-Blundell test rejects the exogeneity of our collaboration variables. Subsequently, we use the Sargan/Hansen's statistic to test that our instruments satisfy the over-identifying restrictions. Notice that, given that some of our instruments were only available from 1993 onwards, the number of observations is reduced in the regressions with instruments.

¹³ The autocorrelation test of the Arellano-Bond estimator rules out that the residuals' dynamic structure is a source of autocorrelation and is thus an ignored cause of bias of the estimates. The Sargan test assumes that the model is identified and tests the validity of the over-identifying restrictions; in our case that the depth of the lags of the dependent and other regressors used as instruments is sufficient to rule out their endogeneity.

5. Empirical Results

In this section we present our estimates of the impact of the degree of industry collaboration. We first introduce our main results and perform robustness checks. Finally, we analyze whether the results differ across observed categories of researchers.

5.1. Main Results

Table 3 reports the basic estimates of research output, measured as the normal *count* of publications. Column 1 displays the estimates of the non-instrumented GLS with fixed effects (GLS FE) model. Column 2 shows the estimates of our benchmark model, the GLS with fixed effects and instrumental variables (GLS FE IV). Column 3 adds *one- and two-year lagged counts* of publications as explanatory variables (GLS FE IV lags).¹⁴ Column 4 uses the Arellano-Bond GMM model, with lagged endogenous and exogenous variables and year dummies as instruments (GMM-AB).

At the bottom of the table, we include *goodness of fit* statistics. For each GLS specification, we report the R^2 and the F-statistic associated with the joint significance of all regressors and the joint significance of the instruments. The null of joint non-significance is rejected in all the models. For the GMM models, we report (i) the Wald χ^2 tests, which reject the joint non-significance of the regressors; (ii) the Sargan/Hansen tests, which are insignificant, suggesting that the models do not suffer from over-identification, and (iii) the Arellano-Bond tests, which do not reject the null that there is an absence of third (or higher) order correlation of the disturbance terms of our specifications, which is required for the consistency of these estimates.

We proceed by reporting the effects of funding and degree of collaboration. Notice that the *had some EPSRC funding* variable allows us to compare the predicted number of publications for any degree of collaboration (including zero) to the predicted number of publications for a researcher without funding. In total, we have a “baseline” productivity prediction, i.e., the expected number of publications for an academic who does not have any funding ($hf=0$ and $fin=0$ in the equation in section 4.1), an additional effect for those that have non-collaborative funding ($hf=1$ and $fin=0$) and another effect emanating from the degree of collaboration ($hf=1$ and $fin>0$). As shown in Table 3, all the funding and degree of collaboration coefficients have the same sign (and are all significant) in all the specifications. Thus, we first explain the results using the benchmark specification. Then

¹⁴ Although the GLS IV estimator does not correct for the autocorrelation created by the endogeneity of lagged publications, we include this specification to compare the resulting coefficients with those obtained using the GMM-AB estimator.

we compare the magnitudes of the coefficients of the benchmark to those of the non-instrumented specification to shed some light on which of the selection mechanisms described in the conceptual framework is consistent with the empirical findings.

Effect of funding and degree of collaboration with industry. The antilog of the constant term minus one, which in the benchmark specification in column two is equal to 0.60, is the baseline productivity prediction, i.e., the expected number of publications for an academic at the lowest rank (lecturer) who does not have any funding or patents.

Turning to our measure of funding, we find that, consistent with our conceptual framework, the positive and significant coefficient for the variable *had some EPSRC funding*_{*t-1*} (β_1 in the equation in section 4.1) corroborates that (non-collaborative) research funding enhances research output. The coefficient implies that for a lecturer with no patents the marginal effect of having non-collaborative EPSRC funding compared to not having any EPSRC funding at all is equal to 0.35 (additional) publications.¹⁵

The linear coefficient of the *fraction of EPSRC funding with industry* variable (β_2 in the equation in section 4.1) is positive and significant (0.925) and the coefficient of the quadratic term (β_3 in the same equation) is negative and significant (-1.710). These results indicate that the effect of the degree of industry collaboration on the number of publications has an inverted U-shape. According to the estimated curve, the *fraction of EPSRC funding with industry* that would result in the maximum number of publications is 0.31.¹⁶ Thus, in the range of 0% to 31%, increasing the fraction of collaborative EPSRC funding results in more publications, but beyond that threshold a higher fraction is associated with a decreasing number of publications. In other words, all else equal, researchers who have approximately one third of their EPSRC funding in collaboration with the industry achieve the highest level of research output.

Figure 2 shows the impact of funding and the degree of industry collaboration on publications for the benchmark specification (GLS FE IV) as well as that of the other specifications (GLS FE, GLS FE IV with lags, and GMM-AB). Using the estimates of each model, we plot the predicted number of publications against the degree of collaborative EPSRC funding for a lecturer with no patents. The degree of collaborative EPSRC funding ranges from 0% to 100%, i.e., from no funding involving industry partners (all

¹⁵ This marginal effect is the difference between the baseline number of publications when the academic had some EPSRC funding - the antilog of $(0.471+0.199)$ minus one - and the baseline publications calculated as described above - the antilog of 0.471 minus one.

¹⁶ This is the antilog (minus one) of the fraction x satisfying the first order condition of the number of publications' maximization problem, i.e., $x^* = \beta_2 / [-2 * \beta_3]$.

non-collaborative) to all funding including industry partners. For the benchmark specification, we also plot the predicted number of publications for a researcher that has not received any EPSRC funding (the predicted number is similar for the other models).

Comparing the GLS IV FE benchmark to the non-instrumented specifications. For all specifications, the intercept and the linear coefficient are positive and significant and the quadratic term is negative and significant. Thus, independently of the estimation method chosen, the effect of the degree of industry collaboration on the number of publications is curvilinear and the curve has an inverted U-shape, as for the benchmark. Noticeably, the curve estimated without instrumenting (GLS FE) is flatter, peaks at higher degrees and, more importantly, lies below the GLS specifications that instrument collaboration (GLS FE IV, GLS FE IV with lags) for degrees of collaboration below 80% and above for those above 80%. Thus, not accounting for the endogeneity problem results in a negative bias for low and medium degrees of collaboration and a positive bias for high degrees.

Relating these empirical findings to our conceptual framework, we conclude that when we instrument the degree of industry collaboration using measures of economic activity around the academics' universities, and thus control for *supply side* opportunities for collaboration, we remove part of the *positive bias* on publications for high degrees of collaboration introduced by the presence of a large proportion of industry-savvy researchers. These academics manage to select and to be selected by high-quality partners. Contrarily, the negative bias of the un-instrumented specification observed at the lower end of the degree of collaboration could be due to the presence of a large proportion of non-industry-savvy researchers that have not developed the same screening capacity, and are not as appealing academic partners to collaborate with.

The negative bias for low degrees of collaboration is also consistent with the existence of less-able academics becoming less output-driven and choosing (or being forced to choose) degrees of collaboration that are too low. In contrast, the negative bias obtained for intermediate degrees of collaboration is inconsistent with the hypothesis that talented and output-driven academics would choose intermediate degrees of collaboration. The positive bias observed for high degrees of collaboration is also inconsistent with the hypothesis that less-able and less output-driven academics choose degrees of collaboration that are too high.

Finally, the potential positive bias of the highly selective academics at the lower end of the distribution of degrees of collaboration does not exist or must be overcome by the negative bias introduced by the non-industry-savvy researchers or the less

-output-driven, i.e., the proportion of highly able selective academics in the lower degrees of collaboration appears to be low. At high degrees of collaboration, the plausible negative bias introduced by lower-able academics undertaking a high number of collaborative projects with low publication potential, seems to be dominated by the positive bias introduced by the industry-savvy researchers.

Effects of past research output. Note that the curve obtained using the GMM-AB specification is above all other curves for all degrees of collaboration. This is because, as shown in column 4 of Table 3, the GMM-AB coefficient associated with the previous year's publications is positive, significant, and large. Because we have taken logarithms, we can interpret this coefficient as an elasticity. Thus, an increase in the number of publications in the previous year by 100% (i.e., doubling them) would result in the following year's expected number of publications increasing by 82 percentage points.

As in earlier papers, our estimates suggest that there is persistence in publications. Several explanations are possible. First, there can be a "Matthew effect" (Merton, 1968), which describes the possibility that the work of those with a higher number of publications receive greater recognition than equivalent work by those that publish less. Second, a higher number of publications reinforces the attractiveness of the academic as a collaborator, increasing her chances of obtaining more and/or better industrial partners and, in turn, higher returns on her collaborations in terms of research output.

Still, the GMM-AB results do not change qualitatively our results and corroborate that the relationship between the degree of collaboration with industry and research output can be represented as an inverted U-shape even when we control for the positive effect of past research performance.

Effects of the other explanatory variables. We control for the *number of patents* filed by the academic to take into account other constraints on the scope of research and/or in the dissemination of research results and to compare our results to previous papers (e.g., Azoulay et al., 2009; Breschi et al., 2008). In accordance with the recent literature, having filed patents in the current year (t) is positively associated with publications, both for the GLS IV FE and the GMM-AB models. The marginal increases, however, are small.¹⁷ In the GLS IV FE model, having filed patents in each of the previous two years ($t-1$ and $t-2$) is also positive and significant. The marginal increases associated with patents filed in

¹⁷ The associated effect is calculated as the antilog (minus one) of the coefficient of the indicator variable and ranges from 0.04 to 0.06 extra publications.

current and in previous years are very similar, thus not suggesting a publication delay. These coefficients are positive but not significant in the GMM-AB model.

We further control for academic rank but we find a significant effect of seniority on publications only for the non-instrumented model. The effect of seniority in the instrumental variables' regressions is absorbed by the instrumented funding variable. Seniority may thus be better at explaining access to funding than publication counts.

5.2. Robustness checks

In Table 4, we reproduce the results of our benchmark model using different measures of research output and collaboration and a balanced sample of academics.

Co-author-weighted count of publications. The first column shows the impact of the fraction of collaborative EPSRC funding when publications are weighted by the number of co-authors. The baseline number of the co-author weighted publication count is 0.33. The coefficients of the linear and quadratic terms of the *fraction of EPSRC funding with industry* variable are significant. The effect of the *had some EPSRC funding* variable is insignificant, suggesting that funding may increase the number of publications simply by increasing the size of the teams. This is further suggested by the negative effect of the professor dummy, as these may primarily benefit from larger labs through co-authorship. The effect of patents is positive and significant as in the benchmark regression in Table 3.

Count of basic and applied publications. Columns 2 and 3 decompose the effects by research orientation. We report the estimates of the impact of the degree of collaboration on the count of applied ("technology") and basic ("science") publications. The baseline number of applied and basic articles is, respectively, 0.33 and 0.11. Thus, the expected number of basic publications is lower than that for applied publications. The existence of EPSRC funding positively impacts the number of basic publications, but the fraction of collaborative EPSRC funding does not. For applied research, the linear and quadratic coefficients associated with industry collaboration variables are instead significant, similar to the benchmark regression in Table 3. This is due to the fact that we consider the field of engineering, where the majority of publications are classified as "applied."

The effect of patents also differs by type of research. Our results in the benchmark specification in Table 3 show that having filed a patent in the current and in each of the two previous years significantly increases the overall number of publications. Columns 2 and 3 in Table 4 show that when separating the effect for applied and basic publications, all the coefficients of having filed patents retain the positive sign. There are interesting

differences, however, in terms of magnitude and significance. We find a significant positive contemporaneous effect of patenting on basic publications and a delayed significant positive effect on applied ones. Indeed, while basic research may produce a variety of complementary research outputs that can result in publications and patents, it may be delaying the publication of more applied, technically-oriented, research papers.

Two-year based funding stock. The estimates in column 4 use a variation of the main explanatory variables. Here, the variables *had some EPSRC funding* and *fraction of EPSRC funding with industry* include the stream of funds received in the last *two years only* (as opposed to the last five). Although this choice is a less accurate reflection of the funding profile of the academic, it deals better with potential autocorrelation issues. Additionally, basing our measure on a shorter window allows younger and more mobile researchers to enter the sample. All coefficients of interest have the same sign and similar magnitude as in the main five-year stock regressions. The degree of collaborative funding resulting in the maximum number of publications is lower than that of the benchmark regression (0.17 as opposed to 0.31). This may be due to the fact that funded projects usually last longer than two years. The positive, long-term effects of collaboration may not be well captured and the degree-maximizing output is smaller. Instead, the model attributes a very large part of the variation in publications to the variable *had some EPSRC funding*.

Balanced sample. The specification in column 5 is estimated using only those researchers that can be observed for the full last 15 years of our sample, so that we are able to build the five-year funding and industry collaboration variables and estimate a balanced 10-year panel. This specification enables us to explore whether the full-sample estimates have been significantly affected by attrition. Again, all coefficients of interest have the same sign and similar magnitude as in the full-sample regressions, which we interpret as being an indication that attrition has not caused important biases in the main estimates. The degree resulting in maximum number of publications is also very similar.

5.3. Results by categories of academics

In Table 5, we test how our benchmark model results differ across categories of academics by interacting our main explanatory variables (*had some EPSRC funding* and *fraction of EPSRC funding with industry*) with different group-indicator variables. For each categorization, the group indicator variable takes the value 0 if the academic belongs to the so-called *reference group* and the value 1 if the academic belongs to the *non-reference group*. Thus, the main coefficient of a given regressor (for instance,

fraction of EPSRC funding with industry_{t-1}) reflects its effect for those in the *reference group*. The effect for those in the *non-reference group* is obtained by adding to the main coefficient the coefficient associated to the interacted term (*group indicator * fraction of EPSRC funding with industry_{t-1}*), if significantly different from zero. As explained below, the trade-off between industry collaboration and publications exists for all the categories of researchers analyzed. That is, the effect of the linear term of the variable *fraction of EPSRC with industry* is positive and that of the quadratic term is negative for all specifications, both for the reference and the non-reference groups. The magnitudes of the effects, however, differ.

The first column distinguishes between academics who are *high collaborators* in terms of having *an above the median share of publications co-authored with industry* (*reference group*) from those that have an average below the median. The main effects for the linear and quadratic terms of the variable *fraction of EPSRC with industry* are larger than those in the benchmark case in Table 3. Thus the relationship between the degree of collaboration and publications for *high-collaborators* is characterized by a more concave curve than the relationship corresponding to the benchmark's estimates. The interaction terms' estimates are significant and of the opposite sign, albeit of a smaller magnitude, than the main effects. This indicates that the relationship for the group of *low-collaborators* is still an inverted U-shape, albeit less concave. Therefore, occasional collaborators also experience gains and losses when varying their degree of collaboration, but the effects are weaker. Indeed, the breadth of ideas, the funding obtained, as well as the constraints imposed, seem relatively more important for heavy collaborators.

The second column distinguishes academics with an above-the-median percentage of *average lifetime of EPSRC funding with industry* (*reference group*) from those below. The magnitude of the main effects on publications is larger than those in the benchmark model and is also larger than those in column 1. This suggests that high collaborators based on lifetime funding exhibit a relationship between the degree of collaboration and publications that is even more sensitive to changes in the degree of collaboration than the relationship of high-collaborators based on joint publications. The correction estimates for low-collaborators as per this measure are not significant.

The third column separates academics with high levels of funding, i.e., those above the median in the *amount of EPSRC funding received in the last 5 years*, from those with lower levels of funding. Again the main coefficients of the funding variables are larger

than those of the baseline model in Table 3. In addition, the correction coefficients for the non-reference group are not significantly different from zero.

In column 4, we distinguish between academics working in a Russell group university from those that do not. We find similar coefficients to those in the benchmark model and no significant difference between researchers in the group of well-known research intensive institutions from those in lesser-known and less-funded institutions. Lower levels of core funding may force smaller institutions to rely relatively more on external grants (Perkmann et al., 2013) and result in higher degrees of engagement with industry (D'Este and Patel, 2005). But, the trade-offs of industry collaboration in terms of academic output appear to be similar for academics at very different kinds of institutions.

Lastly, column 5 distinguishes between academics that are of a lower academic rank from those holding the rank of full professor. The trade-off associated with the degree of collaboration may have been less pronounced for senior academics. Extensive experience and consolidated networks could make the new insights and the additional funds acquired through collaboration relatively less relevant, but at the same time constraints may also be less important. Young researchers, instead, are at a crucial point of their careers, and their research output is expected to be relatively more sensitive to collaboration (Dasgupta and David, 1994). According to our estimates, though, a professor does not experience a significantly different impact of the degree of industry collaboration compared to someone at a lower academic rank.

Overall, we conclude that although there are significant differences, the curvilinear effect of the degree of industry collaboration holds for different categories of academics. We interpret this as evidence of the robustness of our results.

6. Discussion and Conclusion

The effect of research collaboration on publication outcomes has received little attention in the academic literature to date, which has primarily focused on academic patenting and spin-off formations as channels of interaction between science and industry. Many authors have argued, though, that research collaborations, contract research, and consultancy are far more important channels of knowledge transfer. They are, however, more difficult to measure empirically and even more difficult to compare across institutions and time, which may explain why the literature has paid scant attention to these more collaborative forms of university-industry interactions.

This paper uses homogeneous information on the collaborative grants awarded by the EPSRC, by far the most important funding source of research in engineering sciences in the UK, to measure university-industry collaboration over a 20-year period. We show that the effect of collaboration depends on the share of projects undertaken in collaboration with industry, i.e., on the degree of collaboration. Our results indicate that the number of publications increases both with the presence of EPSRC funding and with the fraction of EPSRC funding in collaboration with industry, but only up to a certain point. For degrees of collaboration above 30%–40%, research output declines.

These results confirm the expectations of our conceptual framework which argues that the degree of collaboration-publication relationship could be described by an inverted U-shaped curve. Indeed, the formation of links with the private sector may boost research output because collaboration can provide new ideas and additional funding. But, high degrees of collaboration can also damage research output, as research ideas may then be of lower value, industry may impose non-disclosure clauses or because extensive collaboration could reduce the time to do research and cause attention problems.

Our results may provide an explanation for some of the (apparently mixed) results in the literature, which has effectively focused on linear relationships. On the one hand, they might explain the positive effects documented in studies that investigate forms of collaboration that require little to no direct interaction (e.g., contact through technology transfer offices and trade fairs, as in Hottenrott and Lawson, 2014). Academic patenting can also be viewed as collaboration requiring low levels of interaction (Agrawal and Henderson, 2002). The evidence of a positive relationship between patenting and publications (Azoulay et al., 2009; Breschi et al., 2008; Fabrizio and DiMinin, 2008; Stephan et al., 2007; van Looy et al., 2005) is then consistent with the idea of patenting being on the increasing part of the inverted U-shaped curve.

But, on the other hand, our results can also explain the negative effects documented in other studies. Previous research that investigates the effects of forms of collaboration that require substantial interaction, as for example in academic start-ups, find a negative effect. Toole and Czarnitzki (2010) show that US academics that receive funding to start or join for-profit firms are more productive than their peers, but that they produce fewer publications after receiving the grant. Goldfarb (2008) tracks a sample of 221 university researchers funded by the NASA and concludes that researchers repeatedly funded by the NASA experienced a reduction in academic output. In these cases, the effect is likely to have been that associated to the decreasing part of the inverted U-shaped curve.

Our results are robust to various measures of academic research output and to various subsamples of academics. Nevertheless, some remarks are in order. Basic research output is positively affected by the presence of EPSRC funding but not significantly affected by the degree of industry collaboration. Conversely, for applied publications we find no significant effect of the presence of EPSRC funding but a significant inverted U-shaped effect of the degree of industry collaboration. Indeed, the effects of collaboration identified in the conceptual framework are especially important for applied research. The availability of financial resources, one of the benefits of collaboration, is key for applied research programs.¹⁸ Research ideas arising from collaboration are also more likely to be turned into applied research papers. At the same time, publication constraints should especially affect applied publications as applied research can be published as patents or publications, resulting in potential publication delays (Perkmann and Walsh, 2009).

A key challenge in our analysis is that the observed degrees of collaboration are not exogenous, but are the result of individual and bilateral choices in a two-sided “market” of academics and firms. Our conceptual framework shows that the direction of the bias one might incur, if selection issues are ignored, depends on which mechanism is at work. Our empirical analysis compares the results of our instrumental-variable benchmark specification to those of a non-instrumented regression to shed some light on which mechanism is the most important in our data. We document a negative bias for low degrees of collaboration and a positive bias for high degrees. This is consistent with the presence of a large proportion of “industry-savvy” researchers at the high-end of the range of the degrees of industry collaboration. Indeed, this presence shall introduce a positive bias in publications because industry-savvy researchers end up being matched with more productive firm partners, resulting in more productive research projects.

Our results also bolster empirical evidence from previous surveys and cross-sectional studies on the effects of collaborative research funding on academic output by showing that these results hold for a large longitudinal sample. Even after controlling for endogeneity, we find supportive evidence of the positive impact of the existence of collaboration, as in Gulbrandsen and Smeby (2005). The negative effect of high degrees of collaboration is also consistent with other survey results (Blumenthal et al., 1996) and cross-section empirical evidence (Manjarres-Henriquez et al., 2009).

¹⁸ Financial rewards, however, might also have a positive impact on the production of basic research because basic and applied research efforts are complementary (Thursby et al., 2007) or because they induce a selection of riskier and more basic research programmes (Banal-Estañol and Macho-Stadler, 2010).

We also find a direct, positive effect of patenting on publications, just as in the earlier literature. The contemporaneous effect of patent disclosure is similar to the one of past patents, giving no evidence of a “secrecy” effect. Yet, when we distinguish between types of research, we find a stronger contemporaneous effect of patents on publications in basic science journals. This may explain the positive correlation found in papers that analyze publication and patenting activity of researchers in basic sciences, e.g., life-science (Azoulay et al., 2009; Breschi et al., 2008) and the lack of contemporaneous correlation found in papers that analyze applied sciences, e.g., engineering (Agrawal and Henderson, 2002). We do, however, find a delayed positive effect of patenting for publications in applied journals, suggesting that applied research may suffer from secrecy. Perkmann and Walsh (2009) indeed argue that more applied projects are more likely to be affected by secrecy because of their immediate commercial viability.

In terms of policy or managerial implications, our findings suggest that program interventions encouraging academic researchers to collaborate with industry may be beneficial. A moderate degree of industry collaboration not only facilitates the transfer of knowledge and accelerates the exploitation of new inventions, but it also increases academic research output. Our results point at the disadvantages of academic policies that favor the separation of tasks at the university. If some faculty members focus on research while others perform other activities such as collaboration and development, scientific output might suffer. Indeed, according to our estimates, two academics collaborating moderately (degree of 30–50%) would publish more than a non-collaborating and a fully collaborating one combined (degrees of 0% and 60–100%).

On the other hand, our results also indicate that there are degrees of collaboration that may be *excessive* in the sense of being detrimental in terms of research productivity. Several factors point to an increase in the degrees of industry collaboration in recent times (Stephan, 2012). As the degrees of involvement with industry increase, more academics may find themselves on the decreasing part of the curve. Indeed, academics might start pursuing research lines that no longer result in breakthrough discoveries resulting in fewer publications. In addition, high degrees of collaboration can negatively affect material and data exchange between academics and thus be damaging to the academic community as a whole (Stephan, 2012). But, high degrees of collaboration can also bring gains in terms of patenting or better employment prospects for graduates.

Ours can only be a first step in the analysis of the effects of the various channels of knowledge transfer. We had to limit our analysis to research collaborations sponsored

through public funding, which can only proxy for the extent of the collaboration activities. Including private partners in these grants, though, is highly correlated with obtaining direct funding from the industry (Meissner, 2011). We therefore expect that the curvilinear effect would remain unhindered if, instead of using our measure of degree of industry collaboration, we were to use the overall degree of engagement. In terms of magnitudes, though, the curve may peak at higher degrees of collaboration. First, mechanically, if direct funding is added in the numerator and denominator, the fraction increases. Second, because of the positive correlation, the researchers with one third of their EPSRC funds with industry shall have a substantial amount of direct funding.

With more comprehensive and homogeneous information, we could also make comparisons between the effects of research collaboration and other channels of knowledge transfer such as consultancy and patents. In our sample, research collaborations have a stronger impact than patents. It might also be of interest to tackle interactions between the different channels. We know little about whether collaboration channels complement or substitute each other. Consultancy, for example, might have a positive effect on research output if and only if it is complemented by collaboration in research. Of course, this is only a conjecture and a challenging task for future research.

Acknowledgements

We thank the editor, Ashish Arora, and two anonymous referees for their constructive feedback. We also thank Michael Ben-Gad, Ines Macho-Stadler, Alistair McGuire, Gabriel Montes-Rojas, Markus Perkmann, Dan Peled, Jo Seldeslachts, Anna Torres, seminar participants at City University London, Sheffield, Munich, Ingenio-CSIC (Valencia), DRUID conference (Copenhagen), JEI (Reus), ASSET conference (Istanbul) and MOVE workshop on R&D and Technology Transfer (Barcelona) for helpful comments, and the EPSRC, the NSF and the EPO for providing the data. We thank Kamil Dabrowski and Xenia Radu for their excellent research assistance. The research benefitted from a Pump priming fund at City University and the ESRC research grant RES-000-22-2806. Albert Banal-Estañol acknowledges financial support from the Ministerio de Ciencia y Tecnologia (ECO2010-15052, ECO2013-43011-P and Ramon y Cajal), and Cornelia Lawson support from the European Commission (FP7 #266588 and DIME Doctoral Mobility Fellowship) and a Collegio Carlo Alberto grant. An earlier version of this paper was circulated under the title “The Impact of Industry Collaboration on Research: Evidence from Engineering Academics in the UK”.

References

- Agrawal, A. and R. Henderson, 2002, "Putting patents in context: Exploring knowledge transfer from MIT", *Management Science* 48, 44-60.
- Arellano, M. and S. Bond, 1991, "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies* 58, 277-297.
- Arora, A., P. David and A. Gambardella, 1998, "Reputation and Competence in Publicly Funded Science: Estimating the Effects on Research Group Productivity", *Annales d' Economie et de Statistique* 49-50, 163-198.
- Audretsch, D. and P. Stephan, 1996, "Company-Scientist Locational Links: The Case of Biotechnology", *The American Economic Review* 86, 641-652.
- Azoulay, P., W. Ding and T. Stuart, 2009, "The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research", *The Journal of Industrial Economics*, 57(4), 637-676.
- Baldini, N., 2008, "Negative effects of university patenting: Myths and grounded evidence", *Scientometrics*, 75 (2), 289-311.
- Banal-Estañol, A., and I. Macho-Stadler, 2010, "Scientific and Commercial Incentives in R&D: Research vs. Development?" *Journal of Economics and Management Strategy*, 19 (1), pp. 185-221.
- Banal-Estañol, A., Macho-Stadler, I. and D. Pérez-Castrillo, 2013, "Research Output from University-Industry Collaborative Projects", *Economic Development Quarterly* 27 (1), pp. 71-81.
- Banal-Estañol, A., I. Macho-Stadler and D. Pérez-Castrillo, 2014, "Endogeneous Matching in University-Industry Collaboration: Theory and Empirical Evidence from the UK", working paper Universitat Pompeu Fabra.
- Benavente, J. M., G. Crespi, L. Figal Garone and A. Maffoli, 2012, "The impact of national research funds: A regression discontinuity approach to the Chilean FONDECYT", *Research Policy* 41, 1461-1475.
- Blumenthal, D., N. Causino, E. G. Campbell and K. S. Louis, 1996, "Relationships between Academic Institutions and Industry in the Life Sciences – An Industry Survey.", *New England Journal of Medicine*, 334 (6), pp. 368-73.
- Blumenthal, D., M. Gluck, K. S. Louis, M. A. Stoto and D. Wise, 1986, "University-industry Research Relationships in Biotechnology: Implications for the University", *Science* 13 June, 232, 4756, 1361-1366.
- Blundell, R. and S. Bond. 1998, "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics* 87, 11-143.

- Breschi, S., F. Lissoni and F. Montobbio, 2008, ""University patenting and scientific productivity. A quantitative study of Italian academic inventors", *European Management Review* 5, 91-110.
- Carayol, N., 2003, "Objectives, agreements and matching in science-industry collaborations: reassembling the pieces of the puzzle", *Research Policy* 32, 887-908.
- Cohen, W. M., R. R. Nelson and J. P. Walsh, 2002, ""Links and Impacts: The Influence of Public Research on Industrial R&D", *Management Science* 48, 1-23.
- Czarnitzki, D., C. Grimpe and A. Toole, 2014, ""Delay and Secrecy: Does Industry Sponsorship Jeopardize Disclosure of Academic Research?" *Industrial and Corporate Change*, forthcoming.
- D'Este P. and P. Patel, 2005, "University-industry linkages in the UK: What are the factors underlying the variety of interactions with industry?" *Research Policy* 36, 1295-1313.
- Dasgupta, P. and David, P. A., 1994, "Toward a new economics of science" *Research Policy* 23 (5), 487-521.
- European Commission, 1995, "Green Paper on Innovation", Luxemburg.
- Fabrizio, K. and A. DiMinin, 2008, "Commercializing the Laboratory: Faculty Patenting and the Open Science Environment", *Research Policy* 37, 914-931.
- Florida, R. and W. M. Cohen, 1999, "Engine or Infrastructure? The University Role in Economic Development" In: Branscomb, L.M., Kodama, F., Florida, R. (Eds.), *Industrializing Knowledge: University-Industry Linkages in Japan and the United States*. MIT Press, London, 589-610.
- Geuna, A. and L. Nesta, 2006, "University patenting and its effects on academic research: The emerging European evidence", *Research Policy*, 35(6), 790-807.
- Godin, B., 1996, "Research and the practice of publication in industries", *Research Policy* 25, 587-606.
- Goldfarb, B., 2008, "The Effect of Government Contracting on Academic Research: Does the Source of Funding Affect Scientific Output?", *Research Policy* 37, 41-58.
- Gulbrandsen, M. and J.C. Smeby, 2005, "Industry Funding and University Professors' Research Performance", *Research Policy* 34, 932-950.
- Hall, B., 2004, "University-Industry Research Partnerships in the United States", in Contzen, Jean-Pierre, David Gibson, and Manuel V. Heitor (eds.), *Rethinking Science Systems and Innovation Policies*, Purdue University Press.
- Henderson, R., A. Jaffe and M. Trajtenberg, 1998, "Universities as a Source of Commercial Technology: a Detailed Analysis of University Patenting, 1965-1988", *Review of Economics and Statistics* 80, 119-127.

- Hottenrott, H. and C. Lawson, 2014, "Research Grants, Sources of Ideas and the Effects on Academic Research", *Economics of Innovation and New Technology*, 23 (2), 109-133.
- Jacob, B. A. and L. Lefgren, 2011, "The impact of research grant funding on scientific productivity", *Journal of Public Economics* 95: 1168-1177.
- Laursen, K., T. Reichstein and A. Salter, 2011, "Exploring the Effect of Geographical Distance and University Quality on Industry-University Collaboration in the UK", *Regional Studies* 45, 507-523.
- Laursen, K., and A. Salter, 2004, "Searching low and high: what types of firms use universities as a source of innovation", *Research Policy* 33, 1201-1215
- Laursen, K., and A. Salter, 2006, "Open for innovation: The role of openness in explaining innovation performance among UK manufacturing firms", *Strategic Management Journal* 27, 131 - 150.
- Lawson, C., 2013a, "Academic Patenting: The Importance of Industry Support", *Journal of Technology Transfer* 38 (4), 509-535.
- Lawson, C., 2013b, "Academic inventions outside the university: Investigating patent ownership in the UK", *Industry and Innovation*, 20 (5), 385-398.
- Lee, Y. S., 2000, "The sustainability of university-industry research collaboration: An empirical assessment", *The Journal of Technology Transfer*, 25 (2), 111-133.
- Link A. N. and J. T. Scott, 2005, "Universities as partners in US research joint ventures", *Research Policy* 34, 385-393.
- Lissoni, F., P. Llerena, M. McKelvey and B. Sanditov, 2008, "Academic patenting in Europe: new evidence from the KEINS database", *Research Evaluation* 17, 87-102.
- Manjarres-Henriquez, L., A. Gutierrez-Gracia, A. Carrion-Garcia, and J. Vega-Jurado, 2009, "The effects of university-industry relationships and academic research on scientific performance: Synergy or substitution?" *Research in Higher Education*, 50 (8), 795-811.
- Mansfield, E., 1995, "Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing", *Review of Economics and Statistics* 77, 55-65.
- Maurseth, P., and B. Verspagen, 2002, "Knowledge Spillovers in Europe: A Patent Citations Analysis", *Scandinavian Journal of Economics*, 104(4) 531-45.
- Meissner, C., 2011, "University Research and Industry Involvement. Three Essays on the Effects and Determinants of Industry Collaboration and Commercialisation in Academia", PhD Thesis, City University London.
- Merton, R.K., 1968, "The Matthew Effect in Science: the Reward and Communication Systems of Science Considered", *Science* 159, pp. 56-63.
- Meyer-Krahmer, F., Schmoch, U., 1998, "Science-based technologies university-industry interactions in four fields", *Research Policy* 27, 835-852.

- Mindruta, D., 2013, "Value Creation in University-Firm Research Collaborations: A Matching Approach", *Strategic Management Journal* 34, 644-665.
- Mohnen, P. and C. Hoareau, 2003, "What Type of Enterprise Forges Close Links with Universities and Government Labs? Evidence from CIS 2", *Managerial and Decision Economics* 24, 133-146.
- Mowery, D., R. Nelson, B. Sampat, and 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, 99-119.
- Narin, F., G. Pinski, and H. Gee, 1976, "Structure of the Biomedical Literature," *Journal of the American Society for Information Science*, 25-45.
- Nelson, R., 2004, "The market economy, and the scientific commons", *Research Policy*, 33(3), 455-471.
- Ocasio, W., 1997, "Towards an attention-based view of the firm," *Strategic Management Journal* 18, 187-206.
- OECD (Organisation for Economic Co-operation and Development) 2009. Business-Funded R&D in the Higher Education and Government Sectors, in: *OECD Science, Technology and Industry Scoreboard 2009*. Paris: OECD Publishing.
- Perkmann, M., V. Tartari, M. McKelvey et al., 2013, "Academic engagement and commercialisation: A review of the literature on university-industry relations", *Research Policy* 42, 423-442.
- Perkmann, M. and K. Walsh, 2009, "The two Faces of Collaboration: Impacts of University-Industry Relations on Public Research", *Industrial and Corporate Change* 18, 1033-1065.
- Rosenberg, N., 1998, "Chemical Engineering as a General Purpose Technology", in: Helpman, E. (Ed.), *General Purpose Technologies and Economic Growth*. Cambridge: MIT Press, 167-192.
- Rosenberg, N. and R. Nelson, 1994, "American Universities and Technical Advance in Industry", *Research Policy* 23, 323-348.
- Salter, A. and B. R. Martin, 2001, "The Economic Benefits of Publicly Funded Basic Research: a Critical Review", *Research Policy* 30, 509-532.
- Siegel, D. S., D. Waldman and A. Link, 2003. "Assessing the Impact of Organizational Practices on the Relative Productivity of University Technology Transfer Offices: An Exploratory Study", *Research Policy* 32, 27-48.
- Stephan, P., 1996, "The Economics of Science", *Journal of Economic Literature* 34, 1199-1235.
- Stephan, P., 2012, "How Economics Shapes Science", Cambridge, MA: Harvard University Press.

- Stephan, P., S. Gurmu, A. J. Sumell and G. Black, 2007, "Who's Patenting in the University?" *Economics of Innovation and New Technology* 61, 71-99.
- Stern, S., 2004, "Do scientists pay to be scientists?" *Management Science* 50, 835-853.
- Subramanyam, L, 1983, "Bibliometric studies of research collaboration: A review" *Journal of Information Science* 6, 33-38.
- Thursby, M., J. Thursby and S. Mukherjee, 2007, "Are there Real Effects of Licensing on Academic Research? A Life Cycle View", *Journal of Economic Behavior & Organization* 63(4), 577-598.
- Toole, A., and D. Czarnitzki, 2010, "Commercializing Science: Is there a University 'Brain Drain' from Academic Entrepreneurship?" *Management Science* 56(9), 1599-1614
- Trajtenberg, M., R. Henderson and A. B. Jaffe, 1997, "University Versus Corporate Patents: a Window on the Basicness of Invention", *Economics of Innovation and New Technology* 5 (19), 19-50.
- van Looy B., K. Debackere and J. Callaert, 2006, "Publication and Patent Behaviour of academic researchers: conflicting, reinforcing or merely co-existing", *Research Policy* 35, 596-608.
- Veugelers, R. and B. Cassiman, 2005, "R&D cooperation between firms and universities. Some empirical evidence from Belgian manufacturing", *International Journal of Industrial Organization* 23, 355-379.

Table 1: List of Universities

Russell Group Universities	Number of ID*	Number of Observations	Other Universities	Number of ID	Number of Observations
Birmingham, University of	193	1523	Aberdeen, University of	45	358
Bristol University	82	581	Aston University	62	573
Cambridge, University of	191	1541	Bangor University	30	179
Cardiff, University of	99	802	Brunel University	117	810
Durham, University of	43	308	City University, London	62	556
Edinburgh, University of	92	724	Dundee, University of	49	400
Exeter, University of	40	317	Essex, University of	29	295
Glasgow, University of	108	1033	Hull, University of	37	338
Imperial College London	268	2109	Heriot Watt University	141	1119
Kings College London	48	330	Lancaster, University of	26	230
Leeds, University of	165	1212	Leicester, University of	37	245
Liverpool, University of	102	888	Loughborough, University of	236	1847
Manchester, University of [†]	322	2614	Reading, University of	47	395
Newcastle, University of	143	1188	Salford, University of	99	788
Nottingham, University of	167	1317	Strathclyde, University of	179	1563
Oxford, University of	97	843	Swansea University	94	856
Queen Mary London	82	575			
Queens University, Belfast	101	920			
Sheffield, University of	176	1293			
Southampton, University of	136	1060			
University College London	127	1045			
Warwick, University of	64	621			
York, University of	27	205			

* Researchers can belong to more than one university during their career. Therefore the numbers of IDs do not add up to 3,991, the number of unique individuals in our sample.

Table 2: Descriptive Statistics

Full Sample (33601 obs)	mean	sd	min	max	Correlation with		
					count _{it}	had some funding _{it}	fraction of funding with industry _{it}
Dependent Variables							
count _{it}	1.64	2.68	0	41	1	0.228***	0.125***
co-author weighted count _{it}	0.59	0.93	0	12.26	0.916***	0.208***	0.106***
average impact factor _{it}	0.44	0.72	0	27.07	0.436***	0.167***	0.100***
applied count _{it}	0.9	1.62	0	24	0.778***	0.206***	0.122***
basic count _{it}	0.46	1.52	0	26	0.712***	0.125***	0.0216***
Explanatory Variables							
had some EPSRC funding _{it}	0.67	0.47	0	1	0.228***	1	0.405***
fraction of EPSRC funding with industry _{it}	0.2	0.35	0	1	0.125***	0.405***	1
patent _{it}	0.03	0.18	0	1	0.124***	0.0794***	0.0577***
lecturer _{it}	0.34	0.48	0	1	-0.183***	-0.194***	-0.0998***
senior lecturer _{it}	0.28	0.45	0	1	-0.104***	-0.0511***	-0.0204***
reader _{it}	0.1	0.3	0	1	0.0853***	0.0453***	0.0276***
professor _{it}	0.27	0.45	0	1	0.242***	0.228***	0.109***

Low vs high collaborators in publications [†] (27667 obs)	Low Collaborators (13484 obs)				High Collaborators (14183 obs)				t-test difference
Dependent Variables	mean	sd	min	max	mean	sd	Min	max	
count _{it}	1.43	2.41	0	37	2.22	3.13	0	41	***
Explanatory Variables									
had some EPSRC funding _{it}	0.59	0.49	0	1	0.73	0.44	0	1	***
fraction of EPSRC funding with industry _{it}	0.19	0.34	0	1	0.3	0.38	0	1	***
Low vs high collaborators in funding [‡] (23645 obs)	Low Collaborators (11188 obs)				High Collaborators (12457 obs)				t-test difference
Dependent Variables	mean	sd	min	max	mean	sd	min	max	
count _{it}	1.71	2.62	0	37	2.26	3.2	0	41	***
Explanatory Variables									
had some EPSRC funding _{it}	0.78	0.42	0	1	0.82	0.39	0	1	***
fraction of EPSRC funding with industry _{it}	0.07	0.19	0	1	0.5	0.41	0	1	***
Low vs high recipients of funding ^{‡‡} (28508 obs)	Low Funding (12853 obs)				High Funding (15655 obs)				t-test difference
Dependent Variables	mean	sd	min	max	mean	sd	Min	max	
count _{it}	0.97	1.73	0	23	2.45	3.3	0	41	***
Explanatory Variables									
had some EPSRC funding _{it}	0.38	0.49	0	1	0.89	0.31	0	1	***
fraction of EPSRC funding with industry _{it}	0.11	0.31	0	0	0.36	0.38	0	1	***
Type of institution (28508 obs)	Russell Group (18374 obs)				Non-Russell Group (8408 obs)				t-test difference
Dependent Variables	mean	sd	min	max	mean	sd	min	max	
count _{it}	2.02	3.02	0	41	1.25	2.16	0	33	***
Explanatory Variables									
had some EPSRC funding _{it}	0.68	0.47	0	1	0.62	0.48	0	1	***
fraction of EPSRC funding with industry _{it}	0.24	0.37	0	1	0.24	0.38	0	1	
Academic rank (28508 obs)	Professor (7955 obs)				Not Professor (20553 obs)				t-test difference
Dependent Variables	mean	sd	min	max	mean	sd	Min	max	
count _{it}	2.86	3.66	0	41	1.36	2.25	0	32	***
Explanatory Variables									
had some EPSRC funding _{it}	0.84	0.37	0	1	0.59	0.49	0	1	***
fraction of EPSRC funding with industry _{it}	0.3	0.36	0	1	0.22	0.37	0	1	***

* p < 0.10, ** p < 0.05, *** p < 0.01.

† High (low) collaborators are those who have an *average lifetime collaborative publications with the industry* above (below) the median. Academics with zero publications excluded.

‡ High (low) collaborators are those who have an *average lifetime collaborative EPSRC funding with the industry* above (below) the median. Academics with zero funding excluded.

‡‡ High (low) funding receivers are those who have received *EPSRC funding in the previous 5 years* above (below) the median.

Table 3: Impact of Industry Collaboration on Research Output[‡]

	1	2	3	4
Model	GLS FE	GLS FE IV	GLS FE IV	GMM AB [†]
Dependent variable	count _{it}	count _{it}	count _{it}	count _{it}
Funding and Collaboration:				
had some EPSRC funding _{it-1}	0.044*** (0.013)	0.199** (0.083)	0.203** (0.083)	0.066* (0.035)
fraction of EPSRC funding with industry _{it-1}	0.373*** (0.099)	0.925*** (0.352)	0.897*** (0.344)	0.286* (0.167)
fraction of EPSRC funding with industry _{it-1} ²	-0.421*** (0.145)	-1.710*** (0.655)	-1.651*** (0.638)	-0.432* (0.249)
Patent Filed:				
patent _{it}	0.026 (0.019)	0.040* (0.021)	0.040* (0.021)	0.059** (0.024)
patent _{it-1}	0.014 (0.019)	0.037* (0.020)	0.035* (0.020)	0.014 (0.025)
patent _{it-2}	0.037* (0.021)	0.039* (0.021)	0.038* (0.021)	0.017 (0.026)
Academic Rank:				
senior lecturer _{it-1}	0.048*** (0.014)	0.025 (0.018)	0.022 (0.018)	-0.013 (0.009)
reader _{it-1}	0.093*** (0.022)	-0.002 (0.027)	-0.007 (0.027)	0.001 (0.020)
professor _{it-1}	0.109*** (0.025)	-0.003 (0.032)	-0.009 (0.032)	0.017 (0.019)
Lagged Publications:				
count _{it-1}			0.031*** (0.008)	0.820*** (0.046)
count _{it-2}			0.003 (0.007)	0.042*** (0.015)
Constant	0.479*** (0.014)	0.471*** (0.039)	0.453*** (0.040)	0.087*** (0.022)
Number of observations	33601	28508	28508	26782
Number of ID	3991	3975	3975	3724
Number of instruments	0	3	3	164
F (Joint sig. of instr. in 1st stage)		14.25***	14.25***	
R ² (overall)	0.247	0.087	0.369	
F	19.072***	8.961***	8.709***	
Wald chi2				12597***
AR(1) test z (p-value)				0
AR(2) test z (p-value)				0
AR(3) test z (p-value)				0.7205
Sargan test (p-value)				0.1699

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

‡ The dependent variable, the lagged dependent variables and the fraction of EPSRC funding with industry are in logarithms.

† Endogenous variables are *fraction of EPSRC funding with industry* and lagged *count*. Lagged endogenous and exogenous variables and year dummies are used as instruments.

Table 4: Impact of Industry Collaboration for other Measures of Research Output [‡]

	1	2	3	4	5
Model	GLS FE IV	GLS FE IV	GLS FE IV	GLS FE IV	GLS FE IV
Dependent Variable	co-author weighted count _{it}	applied count _{it}	basic count _{it}	count _{it}	count _{it}
Collaboration stock	5 yrs	5 yrs	5 yrs	2 yrs	5 yrs
Sample	Full	Full	Full	Full	10 yr BP [‡]
Funding and Collaboration:					
had some EPSRC funding _{it-1}	0.079 (0.051)	0.100 (0.074)	0.184*** (0.054)	0.580*** (0.142)	0.243** (0.104)
fraction of EPSRC funding with industry _{it-1}	0.720*** (0.214)	0.952*** (0.307)	0.028 (0.209)	0.992*** (0.331)	0.994** (0.438)
fraction of EPSRC funding with industry ² _{it-1}	-1.535*** (0.395)	-1.785*** (0.575)	-0.509 (0.399)	-3.228*** (0.661)	-2.154*** (0.811)
Patent Filed:					
patent _{it}	0.029** (0.013)	0.030 (0.020)	0.034** (0.015)	0.050** (0.020)	0.043* (0.026)
patent _{it-1}	0.028** (0.013)	0.014 (0.020)	0.012 (0.014)	0.042** (0.020)	0.013 (0.024)
patent _{it-2}	0.026* (0.013)	0.045** (0.020)	0.021 (0.014)	0.043** (0.021)	0.043* (0.026)
Academic Rank:					
senior lecturer _{it-1}	0.012 (0.011)	0.017 (0.016)	-0.004 (0.012)	0.053*** (0.016)	-0.016 (0.023)
reader _{it-1}	-0.003 (0.017)	-0.000 (0.025)	-0.006 (0.018)	0.014 (0.025)	-0.033 (0.033)
professor _{it-1}	-0.034* (0.021)	0.002 (0.030)	-0.010 (0.022)	-0.015 (0.031)	-0.034 (0.039)
Constant	0.282*** (0.024)	0.283*** (0.036)	0.106*** (0.026)	0.236*** (0.071)	0.536*** (0.097)
Number of observations	28508	28508	28508	31837	16250
Number of ID	3975	3975	3975	4436	1625
R ² (overall)	0.001	0.054	0.03	0.131	0.197
F	8.80***	7.93***	2.12***	13.119***	6.986***

Robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[‡] The dependent variable and the *fraction of EPSRC funding with industry* are in logarithms.

^{‡‡} Balanced panel of 10 years.

Table 5: Impact of Industry Collaboration by Groups of Academics[‡]

	1	2	3	4	5
Model	GLS FE IV	GLS FE IV	GLS FE IV	GLS FE IV	GLS FE IV
Dependent Variable	count _{it}	count _{it}	count _{it}	count _{it}	count _{it}
Reference group (group indicator _{it-1} =0)	high collab	high collab	high funding	Russell group	non professor
Non-reference group (group indicator _{it-1} =1)	low collab [†]	low collab [‡]	low funding ^{‡‡}	non Russell group	professor
Funding and Collaboration:					
had some EPSRC funding _{it-1}	0.236** (0.105)	0.142 (0.107)	0.161* (0.087)	0.175** (-0.088)	0.161* (0.087)
group indicator _{it-1} * had some EPSRC funding _{it-1}	-0.090 (0.129)	0.047 (0.139)	0.004 (0.048)	-0.1 (-0.087)	0.259 (0.190)
fraction of EPSRC funding with industry _{it-1}	1.364*** (0.414)	1.580*** (0.430)	1.270*** (0.392)	1.051*** (-0.389)	0.862** (0.382)
group indicator _{it-1} * fraction of EPSRC funding with industry _{it-1}	-0.846** (0.359)	-0.509 (0.389)	-0.073 (0.291)	-0.038 (-0.344)	-0.812 (0.612)
fraction of EPSRC funding with industry ² _{it-1}	-2.687*** (0.814)	-2.723*** (0.845)	-2.177*** (0.743)	-1.386* (-0.74)	-1.677** (0.810)
group indicator _{it-1} * fraction of EPSRC funding with industry ² _{it-1}	1.810** (0.837)	-0.120 (0.902)	-0.685 (0.700)	-0.676 (-0.809)	1.511 (1.212)
Patent Filed:					
patent _{it}	0.040* (0.021)	0.029 (0.021)	0.039* (0.021)	0.040* (-0.021)	0.040* (0.021)
patent _{it-1}	0.036* (0.021)	0.031 (0.021)	0.034* (0.020)	0.037* (-0.02)	0.037* (0.020)
patent _{it-2}	0.040* (0.021)	0.031 (0.022)	0.036* (0.021)	0.038* (-0.021)	0.039* (0.021)
Academic Rank:					
senior lecturer _{it-1}	0.025 (0.019)	0.017 (0.020)	0.022 (0.019)	0.027 (-0.019)	0.030 (0.019)
reader _{it-1}	-0.004 (0.027)	-0.015 (0.029)	-0.012 (0.027)	-0.002 (-0.027)	0.008 (0.028)
professor _{it-1}	-0.006 (0.033)	-0.028 (0.034)	-0.024 (0.033)	-0.004 (-0.032)	-0.091 (0.144)
Constant	0.490*** (0.040)	0.573*** (0.086)	0.518*** (0.042)	0.451*** (-0.079)	0.502*** (0.082)
Number of observations	27667	23645	28508	28508	28508
Number of ID	3833	3150	3975	3975	3975
R ² (overall)	0.105	0.053	0.180	0.101	0.089
F	7.963***	8.209***	9.637***	8.669***	7.923***

Robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

‡ The dependent variable and the *fraction of EPSRC funding with industry* are in logarithms.

† Based on being below the median in the percentage of publications co-authored with the industry. Academics with zero publications excluded.

‡ Based on being below the median in the percentage of average lifetime EPSRC funding with the industry. Academics with zero funding excluded.

‡‡ Based on being below the median in the amount of EPSRC funding received in the past 5 years.

Figures

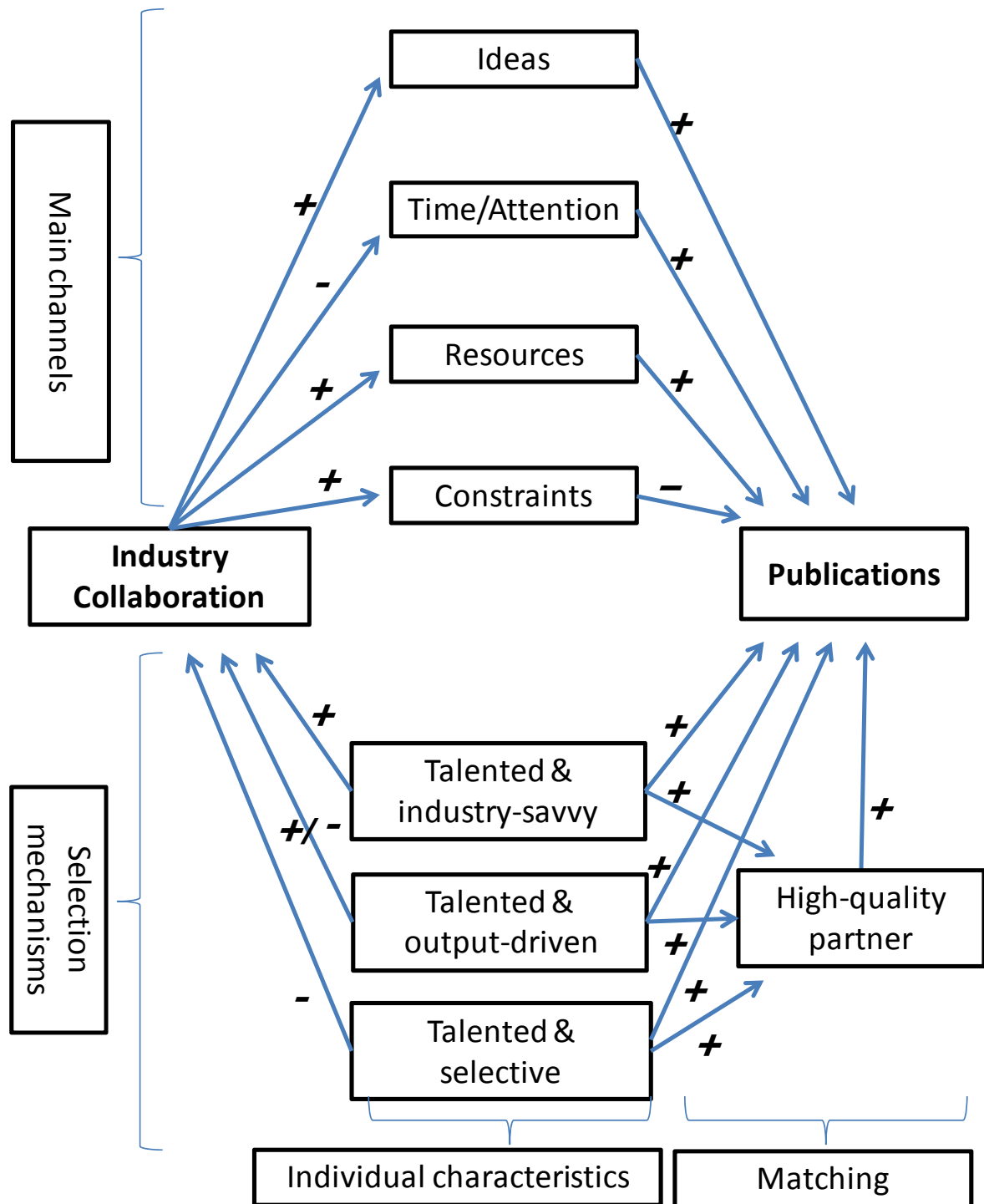


Figure 1. Channels through which collaboration affects/is affected by publications

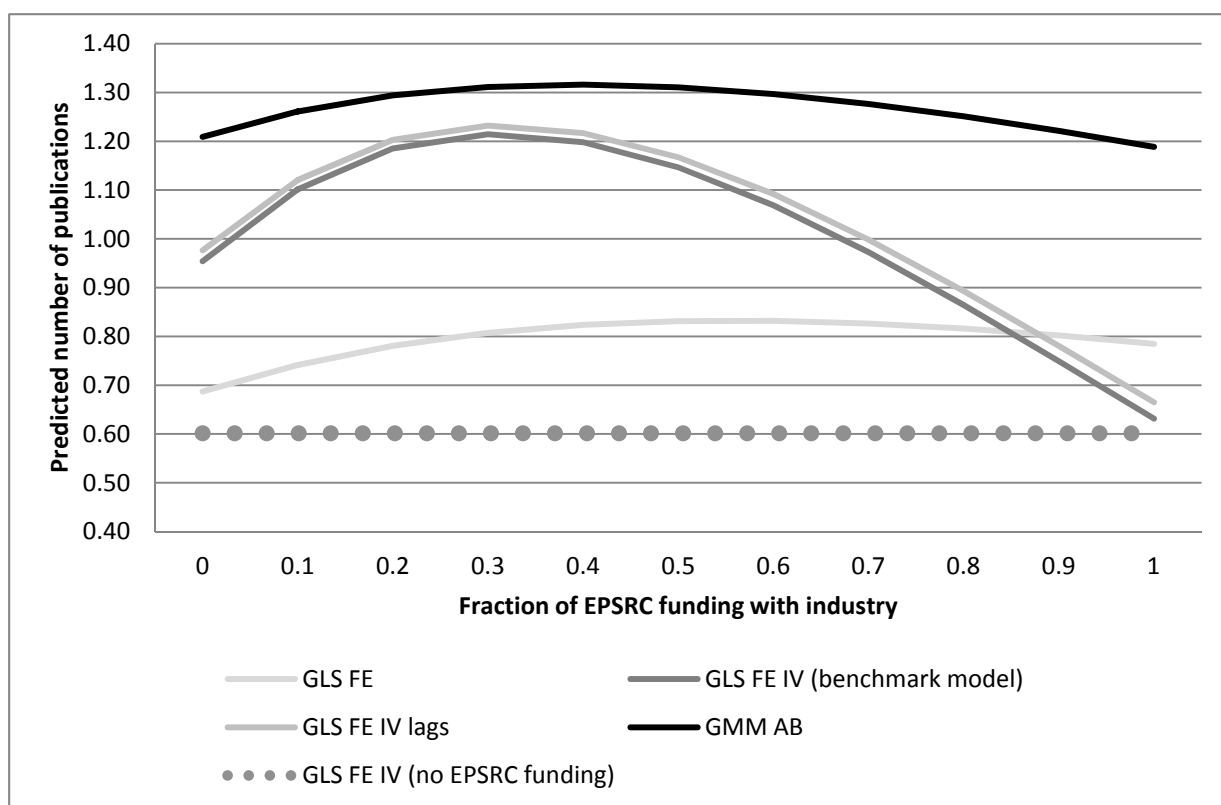


Figure 2: Estimated publications by fraction of EPSRC funding with industry.