GROWINPRO

Growth Welfare Innovation Productivity

Working Paper

Some unintended consequences of commercialisation of basic science: evidence from top scientists and patent-paper-pairs in biosciences

Arianna Martinelli
Alessandro Nuvolari
Elena Romito
Institute of Economics, Scuola Superiore Sant'Anna

7/2020 **March**

This project has received funding from the European Union Horizon 2020 Research and Innovation action under grant agreement No 822781

Some unintended consequences of commercialisation of basic science: evidence from top scientists and patent-paper-pairs in biosciences

Arianna Martinelli[1], Alessandro Nuvolari[1], and Elena Romito[1]

[1]Institute of Economics , Scuola Superiore Sant'Anna, Pisa (Italy)

This paper is produced as part of the GROWINPRO project on Growth Welfare Innovation Productivity that has received funding from the European Union Horizon 2020 research and innovation programme under grant agreement No. 822781.

Abstract

This report aims to investigate the possibility that policies for encouraging industrial exploitation of basic research may hamper basic research itself. We focus on the case of top scientists as they are very active in producing academic knowledge and commercialising their research output. The report contains three main parts. First, we analyse the time evolution characterising the patent activities of top scientists. Second, we produce abundant descriptive evidence on different dynamics characterising academic publishing of *pure scientists* and *scientists-inventors* in the field of biosciences. Third, we carry out patent-paper-pairs analysis to detect the effect of knowledge privatisation on follow-up scientific knowledge in the specific case of cancer research.

Keywords: patents, publications, basic research, top scientists

JEL CODE: O30 O31 O34

1 Introduction

The importance of scientific knowledge for inventive activities in modern economies does not need much elaboration. One of the first works claiming the paramount role of scientific knowledge for more applied stages of the innovation process is the famous report by Vannevar Bush (1945): *Science: the endless frontier*. The author is generally identified within the academic literature as the "father" of the so-called linear model of innovation. This model suggests that the innovation process originates from publicly-funded basic science than flows to applied research and subsequently to production and diffusion. This sequential dimension of the linear model has prompted many critiques as the model, in many instances, appeared as an over-simplistic characterization of innovation processes (Kline, 1985), while neglecting possible non-linearities and feedbacks in the interactions between actors such as universities, firms and government (Etzkowitz and Leydesdorff, 2000). Furthermore, the division of labour between public institutions conducting basic research and private companies focusing only on applied research is not clear-cut. For instance, the search activities of firms, in particular those related to the consolidation of absorptive capacity, have provided significant contributions to scientific developments (Fleming and Sorenson, 2004; Gambardella et al., 1995; Cohen and Levinthal, 1989). These activities represent an important link between corporate science and the scientific community (Cockburn and Henderson, 1998; Rosenberg, 1990). Remarkably, some recent evidence suggest that the relevance of corporate science for scientific development has been decreasing over time (Arora et al., 2015).

While the linear model of innovation might be an oversimplification of the innovative process, it has the merit to place a significant emphasis on the role of scientific development in the innovative process (Balconi et al., 2010). This primary role has weakened over time, in favour of more emphasis on the transmission mechanisms from basic scientific development to commercialisation. Since the 1970s, there has been an increasing tendency of criticising governments policies as not sufficiently effective in promoting the dissemination of publicly funded research to actors who could make use of it. Contextually, there was concern on universities having little incentives to seek practical uses for inventions financed by the governments [1].

The U.S. federal government reacted to this concern in 1980 by establishing the Bayh-Dole Act to allow universities and other publicly funded institutions to patent and exclusively license the outcome of their research. There is consensus within the literature that the Bayh-Dole Act fostered academic patenting and licensing firstly in the U.S. and subsequently, by means of emulation, in several European countries (Shane, 2004a; Mowery et al., 2001). While in Europe there is not a single law comparable to the Bayh-Dole act in the U.S, several national laws *de facto* establish similar regimes in Europe[2]. Because of the budgetary cuts that most universities experienced in the 1990s, these policy changes tremendously impacted on the incentives to commercialise research carried out within universities (Siegel et al., 2003).

In general terms, the process of science commercialisation implies the embodiment of basic knowledge into marketable applications. Together with patenting and licensing, other vehicles for science commercialisation are, for example, start-up companies employing university faculty (i.e. university spin-off companies) and "soft technology transfer" such as research contracts or consulting (Grimpe and Fier, 2010).

These changes have stimulated a growing scientific interest in studying both the determinants and potential impacts on society of this increasing commercialisation of science. As regards the determinants of more involvement of university faculty in commercialisation activities, the research has highlighted several key factors. For instance, individual characteristics play a crucial role as scholars with a high degree of entrepreneurial capacity will be more involved in commercialisation activity either collaborating with existing ventures or establishing their new business (Clarysse et al., 2011). Second, the efficiency of universities Technology Transfer Offices (TTOs) is among the most prominent promoters of academic entrepreneurial activities. In fact, *TTOs serve as an 'intermediary' between suppliers of innovations (university scientists) and those who can potentially (help to) commercialise them, i.e. firms, entrepreneurs, and venture capitalists* (Siegel et al. (2007),

[1]See "GAO Reports and Comptroller General Decisions: Transferring federal technology Administration of the Bayh-Dole Act by research universities" online available at `https://www.gao.gov/assets/230/225671.pdf`

[2]For an overview, see Audretsch and Göktepe-Hultén (2015).

p. 641). Third, also the location of universities matters as the probability of a venture capitalist to invest in a new technology sharply declines with geographical distance (Sorenson and Stuart, 2001). Fourth, the presence of *star scientists* within universities allows overcoming a geographical distance liability to venture capital as they signal research quality within the institution (Fuller and Rothaermel, 2012). The notion of *star scientist* was introduced by Zucker and Darby's seminal works (Zucker and Darby, 1998, 1997, 2001), where the authors call star scientists the most productive scholars into the field of genetic research. In line with Fuller and Rothaermel (2012), the copious Zucker and Darby's research activity indicates that links between academic knowledge and commercial applications considerably increases in presence of star scientists. Following this point, star scientists are an interesting group to focus on when interested in detecting newly emerging trends in the generation of scientific knowledge and its possible commercialisation.

As regards the effect of more involvement of university faculty in commercialisation activities, many scholars agree that the phenomenon is associated with an improved level of consumer welfare as the society can benefit from the output of basic research only in case this is converted into marketable products or services (Bornmann, 2013; Shane, 2004b). Nevertheless, concerns arise about the possibility of distortion of the direction of basic science (Shane, 2004b). First, universities can deviate from their original mission accepting compromises with basic academic values, tempted by the possibility of profiting from their research. This could trigger the occurrence of conflicts of interest (Bok, 2009). Second, the introduction of an economic reward (in addition to the wage) may lead to a distortion in terms of scientists' incentive to innovate. Several studies show that extrinsic rewards can undermine intrinsic motivations traditionally stimulating a taste for free scientific science (Deci et al., 2001; Stern, 2004; Murdock, 2002). Third, while public and proprietary knowledge were characterised by different means of appropriability, the current tendency of making basic knowledge closer to the market also leads to an overlap of appropriability mechanisms. The main concern is the likelihood of reducing access to basic scientific research when this is protected by formal IPRs leading to a reduction in follow-up scientific research that relays on recombination and exploitation of exiting pieces of knowledge (David, 2004; David et al., 2000; Heller and Eisenberg, 1998; Murray and Stern, 2007).

This report seeks to assess how the generation of basic scientific knowledge has been affected by the emergence and consolidation of a new policy regime with a growing emphasis on the commercialisation of research findings of universities and other public research organisations. While most of the literature focuses on the direct effect of science commercialisation on consumer welfare (Bornmann, 2013; Shane, 2004b), we underline potential drawbacks of this tendency on top-level basic research. Given the undeniable importance of basic research for practical applications, a negative result would hint to a possible long-term detrimental effect on social welfare, albeit an indirect one.

1.1 Structure of the report

We analyse this issue using different sources of data to deliver both descriptive and econometric evidence. This report has three main sections.

Section 2 analyses the trends characterising patent activities of top scientists. As already explained, top scientists are the group of academics that simultaneously contribute the most to both basic research and to the process of its commercialisation. We match data on top 1% scientists (i.e. Highly Cited Researchers (HCR)) released by Clarivate Analytics and patent data (retrieved from the EPO-PATSTAT Database) to obtain useful information on top scientists patent activity.

In section 3, we restrict the analysis to the field of biology and biochemistry as these are among the areas that mostly exploited the practice of protecting pieces of basic research with patents. After identifying top scientists who also appear as patents' inventors, we retrieve all the useful information about their academic publishing from the ISI Web of Science Core Collection database, and we produce abundant descriptive evidence on different dynamics characterising *pure scientists* and *scientists-inventors*.

Finally, in section 4 we focus our attention on the effect of knowledge privatisation on follow-up scientific knowledge in the field of *cancer research*. We use patent-paper-pairs to precisely identify pieces of knowledge disclosed by both scientific publications and patents. Then we set up an econometric model, as firstly proposed by Murray (2002), to isolate the effect of the patent's award on the rate of the follow-up stream of knowledge generated by the original paper.

2 Top scientists' patenting trends

This section provides descriptive evidence on the dynamics of patenting and publishing activities carried out buy most productive scientist (i.e. the top scientists) in 2018. The role of market mechanisms in shaping the research agenda of universities and other public research organizations has drastically increased starting from the 1980s', and top scientists are primarily involved in this paradigmatic change of conceiving basic research.

2.1 Data construction

Clarivate Analytics[3] yearly publishes a list of the most influential authors across different scientific fields. These authors are the ones whose scientific production has received the highest number of forward citations (top 1% per each scientific field) within the year. For this analysis, we consider the top scientists listed in 2018 Highly Cited Researchers report (HCR2018) released by Clarivate Analytics. The report includes 6,021 top scientists working on 22 distinct research fields. Table 1 shows the list of all the scientific fields included in HCR2018, the number of top scientists across fields and the share over the total for each category. The shares vary between 1.49% (*Mathematics*) and 4.32% (*Chemistry*), with the exception of *Cross-Field* which includes scientists whose contribution is not circumscribed within a unique field.

Table 2 and table 3 shows the top 5 countries and affiliations of these top scientists. The largest share (43.4%) of top scientists are located in the United States, followed by the United Kingdom, China, Germany and Australia, all with a percentage below the 10%. Interestingly, not all the top 5 affiliations are located in the United States. Table 3 shows that Harvard Universities and Stanford Universities are the two institutions with the highest number of affiliated top scientists, followed by the Chinese Academy of Science, the Max Planck Institute, and the University of Berkeley.

To study the patenting activity of these top scientists, we match their names to the inventor names contained in the EPO-PATSTAT database (Version Autumn 2019)[4]. From this database,

[3]Clarivate Analytics was formerly the Intellectual Property and Science division of Thomson Reuters. This company commercialises the Web of Science database used in this report.

[4]We perform the matching using the Stata function *reclink* and we adopt a conservative approach

Table 1: Scientists distribution across fields

Scientific Field	Num of Scientists	Share
Agricultural Sciences	158	2.62
Biology & Biochemistry	231	3.84
Chemistry	260	4.32
Clinical Medicine	490	8.14
Computer Science	97	1.61
Cross-Field	2007	33.33
Economics & Business	96	1.59
Engineering	204	3.39
Environment/Ecology	185	3.07
Geosciences	183	3.04
Immunology	146	2.42
Materials Science	207	3.44
Mathematics	90	1.49
Microbiology	147	2.44
Molecular Biology & Genetics	248	4.12
Neuroscience & Behavior	195	3.24
Pharmacology & Toxicology	161	2.67
Physics	210	3.49
Plant & Animal Science	221	3.67
Psychiatry/Psychology	154	2.56
Social Sciences. general	210	3.49
Space Science	121	2.01
Total	6021	100

Table 2: Top 5 Countries

Country	Num of scientists	Share
United States	2611	43.36
United Kingdom	546	9.06
China Mainland	481	7.9
Germany	358	5.9
Australia	244	4.05

Table 3: Top 5 Universities

Affiliation	Num of scientists
Harvard Univ	186
Stanford Univ	100
Chinese Acad Sci	91
Max Planck Soc	78
Univ Calif Berkeley	64

we retrieve all the granted patents at the United States Patents and Trademark Office (USPTO).

2.1.1 Descriptive analysis

Overall, we find that almost half of the top-scientists (49.5%) is a *scientists-inventors* as it has at least one patent. On average, our *scientists-inventors* have 11.9 patents, whereas the median is 4.

Figure 1 shows the evolution of the number of patents invented by at least one top-scientist between 1985 and 2015[5]. The graph indicates an overall increase in the patenting activity of scholars that are considered the best performers in conducting basic research. However, we can notice that during the first decade of the 2000s, the number of patents is rather stable.

Figure 1: Time evolution of top scientists patenting

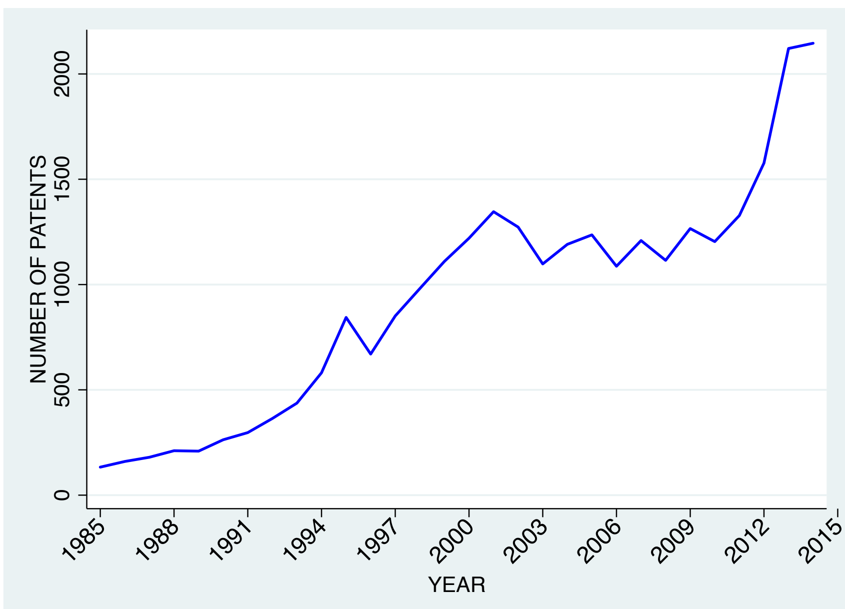


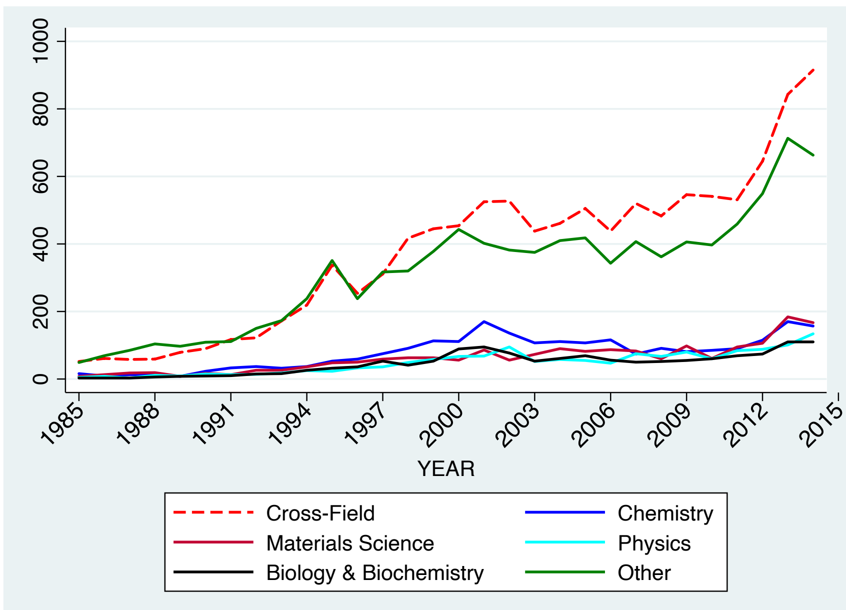
Figure 2 shows a similar graph for the four categories that mostly exploit the practice of

requiring the score of matching is at least 99%.

[5]We include data only up to 2015 to avoid data truncation as we use granted patents. Since the patent lag grant is on average of 4.5 years, including more recent periods would lead to a structural distortion.

patenting basic research. In accordance with the literature, our data show higher rates of academic patenting for *Physics, Material Science, Chemistry* and *bio-sciences (Biology & Biochemistry)*. The graph suggests that these categories display similar patterns, both as regards the magnitude of the phenomena and trends. The graph also reports the time evolution for two other categories which have the most significant number of patents. The first one is the *Cross-field* category that refers to *scientists-inventors* whose scientific contribution is not circumscribed within a unique field. The second one is the residual one (*Other*) which includes all patents of the remaining 17 categories.

Figure 2: Time evolution of top scientists patenting across fields



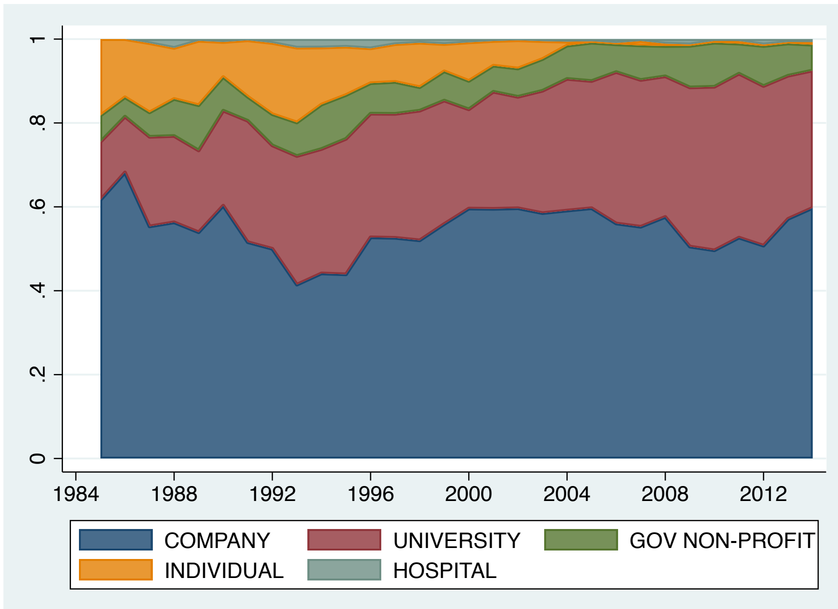
Using the assignee classification provided by the EPO-PATSTAT Database, we can distinguish different categories such as private companies, universities, government non-profit institutions, individuals and hospitals[6].

Figure 3 illustrates the time evolution of the share of each assignee type over time which shows a

[6]For details about this classification see the official EPO-PATSTAT documentation available at `http://documents.epo.org/projects/babylon/eponot.nsf/0/11CE75EDDF734288C125848F0048F533/$FILE/data_catalog_global_v5.14_autumn_2019_en.pdf`

substantial increase of universities as patent assignees. This trend is consistent with the increasing incentives universities have to try to profit from their scientific research directly.

Figure 3: Share of Assignee type

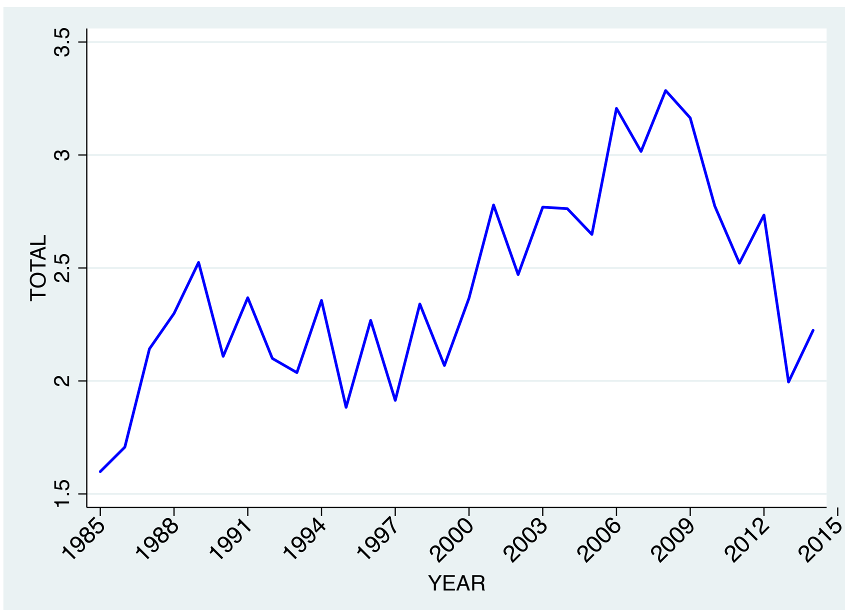


Finally, the innovation literature has used the citation to non patent literature (NPL) as indicators of the contribution of public science to industrial technology (Narin et al., 1997). These citations to scientific articles measure the proximity of an invention to scientific knowledge (Callaert et al., 2006) and to more complex and fundamental knowledge (Cassiman et al., 2008). We retrieve the number of NPL citations for our set of USPTO granted patents from the OECD Patent Quality Database (version 2020) (Squicciarini et al., 2013).

Figure 4 reports the evolution of the normalised[7] number of NPL references reported in *scientists-inventors*' patents. The graph shows an increasing trend up to 2009, followed by a sharp decrease indicating a shift towards less basic and more commercial innovations.

[7]The normalisation is carried out by considering patents in the same technological field and filed in the same year.

Figure 4: Evolution of Non Patent Literature (NPL) over time



3 Scientists with and without patents

In this section, we restrict the analysis on the top scientists in the fields of biology and biochemistry as these areas mostly exploited the practice of patenting basic research (Holman and Munzer, 2000). We investigate whether different dynamics between *pure scientists* and *scientists-inventors* publishing activity emerge.

Furthermore, we add a temporal dimension to the picture comparing the same top-scientists variables across two periods: before and after the first patent application year for each scholar. This analysis further contributes to highlighting how a close link between basic and commercial science can reshape the dynamics of basic research and influence on scientists attitude and behaviours.

3.1 Data

We retrieve from the ISI Web Core Collection Database all the scientific articles published by all the top scientists in the biology and biochemistry category listed in the HCR2018. We use this information to build our variables of interest. First, we measure the scientific productivity of each scholars counting the number of articles published from 1985 up to 2019. Second, we proxy their quality employing a standard measure which is the citations received from each publication (Garfield, 1979; Price, 1965; Posner, 2000), Third, we leverage on the information on articles' access status released by the ISI Web Core Collection Database, to construct an original index to measure scholar's inclination to disclose research output to the rest of academic community freely. Each article is classified within one of the following categories depending on the accessibility status:

- Gold access: the author publishes in an online fully open access journal, and a license for reuse and distribution is implicitly provided;
- Green access: the author publishes into a non-open access journal, but then he self-archives a copy in a freely accessible institutional or specialist online archive known as a repository, or on a website. Also in this case, articles are accompanied by a license;
- Bronze access: the author publishes into a non-open access journal and he self-archives a

freely available copy of the paper into an institutional or personal repository. However, differently from the Green access case, Bronze open access does not provide a license. That means the publication is free to read, but it cannot be reused and distributed (for example in presentations or course material);

- Close access: the author publishes in a non-open access journal, and no freely readable copies are available.

Aggregating all these information at scientist level, we can proxy their attitude towards a more open scientific paradigm.

3.2 Results

The number of top scientists in the *Biology & biochemistry* scientific field is 231 and 116 of them have invented at least one patent. The average number of patents for *scientists-inventors* belonging to this category is 11.7. The average number of publications, instead, is 369 in case of *scientists-inventors* and 107.6 in case of *pure scientists*.

Figure 5 shows the total number of publications by each *pure scientist* on the left, and each *scientists-inventor* on the right. The figure depicts that scientists who also patent their inventions publish relative more than scholars who have never patented. While this is in line with most of the research done (Zucker and Darby, 2001; Fuller and Rothaermel, 2012), results on the quality of publications differ. In fact, Figure 6 shows that papers' average quality is comparable across *pure scientists* and *scientists-inventors*.

This descriptive evidence partially contradicts Torero et al. (2001) who predict that the probability that the scientist conducts joint research with a firm increases as the quality of an academic star bio-scientist increases.

To measure top bio-scientists propensity to share their scientific results freely, we use data on the access status of top bio-scientists publications. For each scientist, we build an accessibility index computed as an arithmetic average of the value (from 1 to 4) assigned to each article depending on its accessibility status. The highest value 4 is assigned to publications with Gold access, whereas

Figure 5: Total publications per top scientist

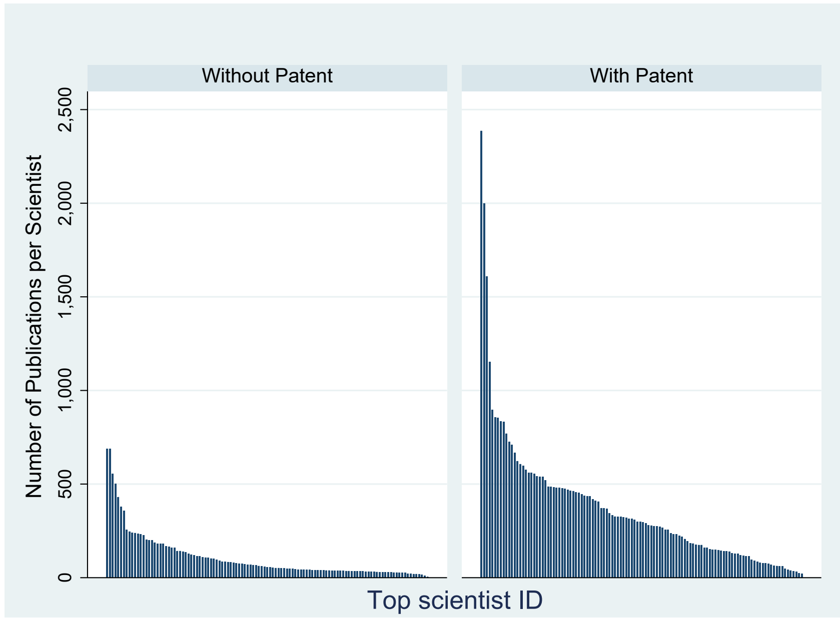
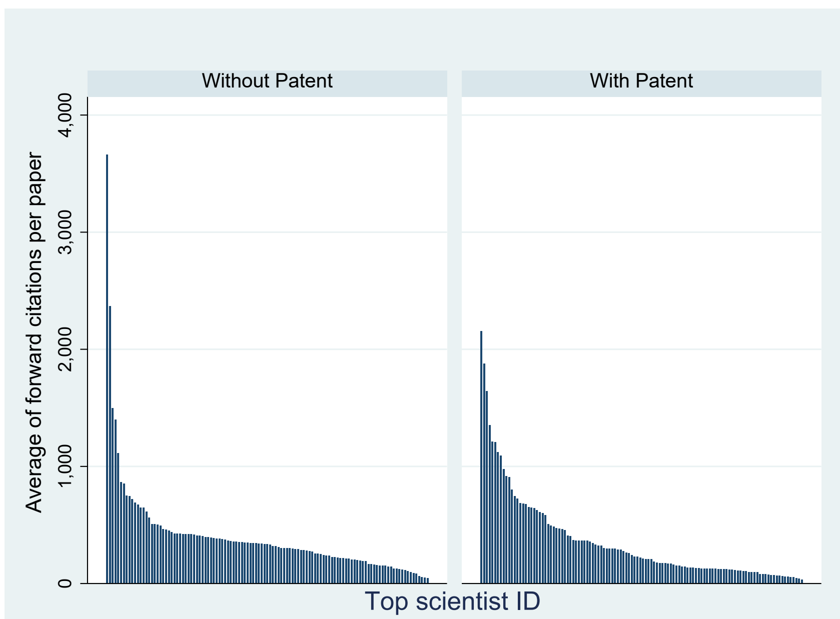


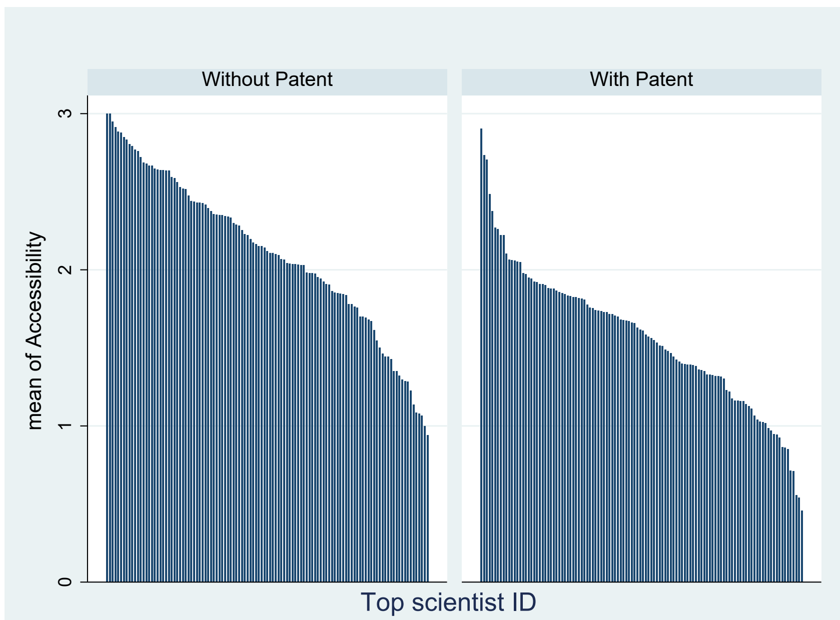
Figure 6: Average citations per publication



the lowest value 1 is assigned to publications with Close access. The resulting indicator ranges between 1 and 4 with higher measures indicating scientists more prone to disclose their research outcome freely[8].

Figure 7 provides an overview on the accessibility index of *pure scientists* and *scientists-inventors*. The figure clearly shows that the index is systematically lower in the case of authors with at least one patent. This difference indicates that pure scientists not only do not privatize their research through patents but they freely disclose more to the rest of the academic community.

Figure 7: Accessibility index per scientist



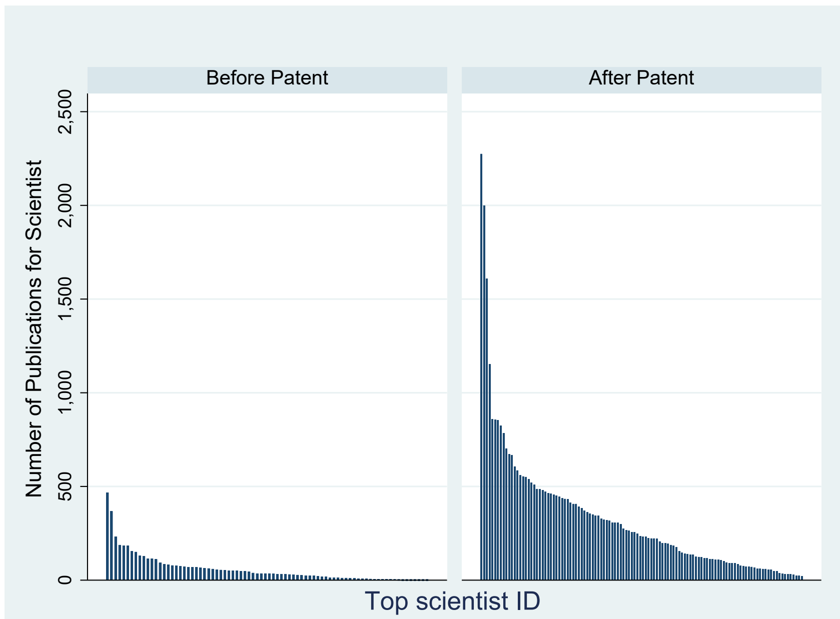
This final part examines the time dynamics of the same variables concerning the granting of a patent. As we are interested in assessing what happens after the patent award, we restrict our sample on the 116 *scientists-inventors*.

Figure 8 shows the total number of articles published before (on the left) and after the first patent grant (on the right). The higher number of publications after the patent might lead the

[8]For instance, if a scientist publishes 10 articles, of which 3 are Closed access (value 1), 4 are bronze access (value 2), 2 are green access (value 3) and 1 is gold access (value 4). The resulting index for this scientist is calculated as: $Accessibility\_index = \frac{(3*1)+(4*2)+(2*3)+(1*4)}{10} = 2.1$

reader to infer that a patent application boosts academic productivity. However, no control, such as scientists' age and period of activity, is included. For instance, the sharp increase observable in the right panel of Figure 8 may simply be due to a longer career after the post-grant period (i.e. more extended period in which cumulate publications).

Figure 8: Total publications per top scientist before and after patent



A different explanation, instead, can be put forward for the quality of publications. As this measure is computed as the average number of citations received by each scientist, it is not affected by a scholar's age and length of the career after the patent. The average number of forward citations received by each scientist is showed in Figure 9. No substantial differences emerge, indicating that the quality of research does not increase once the scholar begins to make his innovation closer to the market.

Finally, as regards accessibility, Figure 10 shows that the propensity to freely disclose academic production increases after patent application. While Figure 7 showed *scientists-inventors* being less prone to publish their articles openly then *pure scientists* colleagues, this new evidence suggests that the source of such a behavioural difference resides within the pre-patent period when scientist might want to hide relatively more research outcome.

Figure 9: Average citations per publication before and after patent

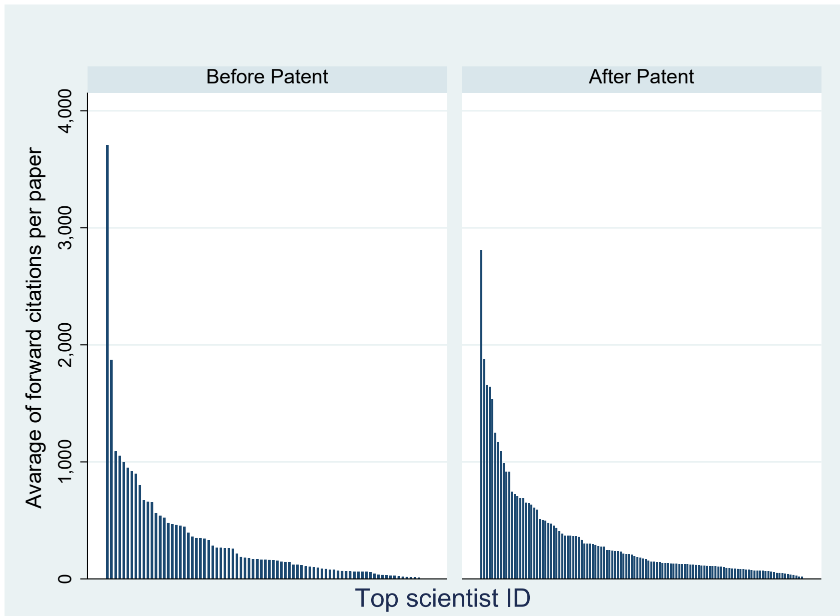
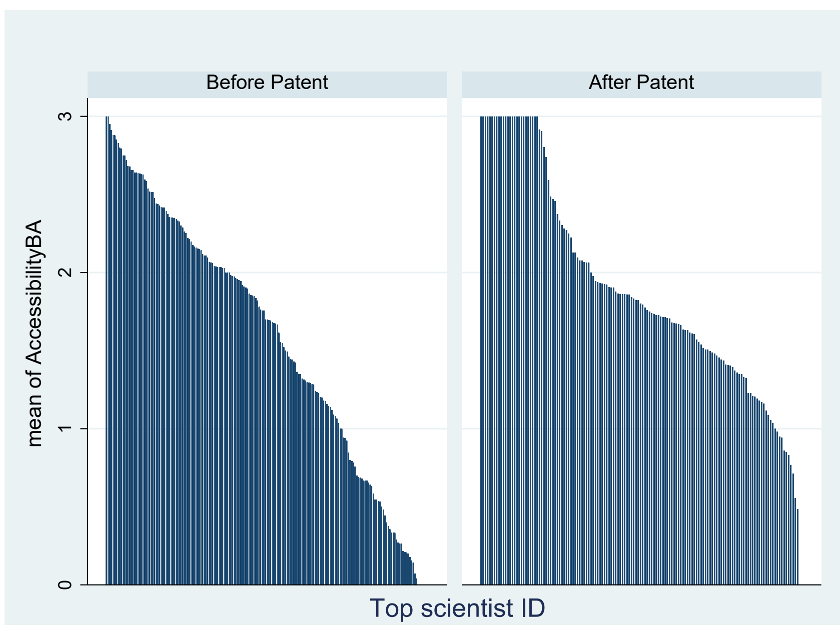


Figure 10: Accessibility index per scientist



4 Patent-Paper-Pairs

In this section of the report, we turn our attention to a specific case of bio-science, namely *research on cancer detection* to provide econometric evidence on the effect of patenting on subsequent scientific development. In particular, we assess whether using intellectual property rights (IPRs) to protect an innovation disclosed initially on a public platform promotes or hampers future cumulative research activities. Furthermore, we focus on two specific aspects. The first one is whether the granting of a patent differently affect the research in the same area carried out by different institutions such as private companies or universities. The second one is whether different actors might be differently affected by the granting of a patent depending on the degree of applicability (i.e. closeness-to-the market) of the patented innovation.

The interest for studying the effect of patent granting on the rate of follow-up scientific research originates from the evidence that while the legal research exemption[9] allows researchers to use proprietary inventions without infringing the monopolistic rights of the patent holder, its real applicability and scope are quite limited (Dreyfuss, 2004). These limitations have been revealed in certain high profile decisions such as, for example, the *Roche Products Inc v. Bolar* and the *Madey v. Duke University.* In the first case, the court narrowed the scope of the exemption hugely indicating that it was limited to experiments "for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry" and did not extend to use for business reasons. In the second case, the court specified further that "the profit or non-profit status of the user is not determinative". The consequence of these decisions is that scientists and universities could be sued for patent infringement if they use proprietary technologies in their research and it may lead to a negative effect on the degree of follow-up scientific research in the same domain. For the scientists not to be suited, they need to negotiate formal access via licensing agreements. The lower the cost of such negotiation, the smaller the potential negative effect on follow-up production of knowledge should be. It is reasonable to think that this transaction cost is different between private and public institutions. While the first has traditionally been involved into the cross licence market and they

[9] See Dent et al. (2006) for a review on research exemptions in different OECD legislations.

operate “at least somewhat efficiently” in minimizing costs (Sampat and Williams, 2019), public actors may lack of some organizational capabilities needed in engaging in the complex licensing process (e.g. different budget, less capable university technology transfer offices).

The questions we try to address are:

- **RQ1** Which is the effect of granting a patent on further scientific research carried out by public and private institutions?
- **RQ2** Does this effect depend on whether the protected scientific discoveries is applied?

4.1 Methodology, variables and data

To answer the two questions, we build an original dataset of PPP (i.e. patents and publications with common content) as firstly suggested by Murray (2002). All the granted patents by the USPTO between 2004 and 2011 in the technological class related to “Detecting cancer” (class 435/6.14) were manually matched to publications included in the WoS (Web of Science) based on correspondence between inventors and authors, patent application dates and publication period, patent and paper abstracts. The starting patent sample includes 1,652 patents, of which 373 are paired to a publication and constitute the set of *paired-patents* used for the empirical analysis.

Following the identification strategy in Huang and Murray (2009) we can treat the time of patent granting as an exogenous shock which allows us to measure the effect of patent granting on the production of subsequent new knowledge (measured as yearly citations to the paired publication) and to isolate the impact of making *proprietary* an innovation that previously was *public*[10]. A reaction in the number of annual citations received by the publication after the granting of the corresponding patents indicates a change in the rate of follow-on public knowledge.

The econometric models we estimate to answer our first research question (**RQ1**) are:

[10]See Appendix A.1 for detailed explanation of empirical strategy and econometric model.

$$CIT\_PUBLIC_{i,t} = \alpha + \beta_0 PATENT\_WINDOW_{i,t} + \\ + \beta_1 PATENT\_IN\_FORCE_{i,t} + \gamma_i + \delta_t + \epsilon_{it} \tag{1}$$

$$CIT\_PRIVATE_{i,t} = \alpha + \beta_0 PATENT\_WINDOW_{i,t} + \\ + \beta_1 PATENT\_IN\_FORCE_{i,t} + \gamma_i + \delta_t + \epsilon_{it} \tag{2}$$

where $CIT\_PUBLIC_{i,t}$ is the number of citations received by the paired publication i in year t from publications whose authors' affiliation is a public research center, and $CIT\_PRIVATE_{i,t}$ is the number of citations received by paired publication i in year t from publications whose authors' affiliation is a firm. The variable $PATENT\_IN\_FORCE_{i,t}$ is a dummy variable that is equal to 1 for all the years t when the patent associated to the publication i is valid (i.e. in force). $PATENT\_WINDOW_{i,t}$ is a dummy variable that is equal to 1 for the year t in which the patent associated to publication i was granted. Finally, (γ_i) is the set of paired publication fixed effects, (δ_t) is a set of year fixed effects, and (ϵ_{it}) is the error term.

To answer the second research question (**RQ2**), we add to the previous models an interaction term, between our variable of interest $PATENT\_IN\_FORCE_{i,t}$ and the dummy variable $APPLIED\_KNOWLEDGE_i$. This variable takes the value 1 if the paired paper i is published in an applied journal according to the CHI-classification (Hamilton, 2003). In a nutshell, the models we are going to estimate are:

$$CIT\_PUBLIC_{i,t} = \alpha + \beta_0 PATENT\_WINDOW_{i,t} + \\ + \beta_1 PATENT\_IN\_FORCE_{i,t} + \beta_2 APPLIED\_KNOWLEDGE_i + \\ + \beta_3 APPLIED\_KNOWLEDGE_i * PATENT\_IN\_FORCE_{i,t} + \\ + \gamma_i + \delta_t + \epsilon_{it} \tag{3}$$

$$
\begin{aligned}
CIT\_PRIVATE_{i,t} = \alpha + \beta_0 PATENT\_WINDOW_{i,t} + \\
+\beta_1 PATENT\_IN\_FORCE_{i,t} + \beta_2 APPLIED\_KNOWLEDGE_i + \\
+\beta_3 APPLIED\_KNOWLEDGE_i * PATENT\_IN\_FORCE_{i,t} + \\
+\gamma_i + \delta_t + \epsilon_{it}
\end{aligned}
\tag{4}
$$

While β_1 represents the effect of an enforceable patent; β_3 represents the additional effect related to the degree of knowledge applicability.

See Appendix A.2 for the descriptive statistics of all the variables.

4.2 Results

Table 4 reports the estimation results related to our first research question (**RQ1**). The coefficients of the variable *PATENT_IN_FORCE* reported in column 1 and 2[11] are negative and significant, indicating that the award of the patent has a negative and significant effect on the rate of follow-on scientific research by a public institution. In particular, we find a decline in the expected number of annual citations of about 2.27. Since the average number of citations is 9.657, this corresponds to a reduction of 23.5% (significant at the 0.01% level). This result confirms the intuition that "privatization" of formerly freely available knowledge stifles further scientific contribution from public actors. The coefficients of the variable *PATENT_IN_FORCE* reported in column 3 and 4 are negative but not statistically significant, indicating that scientific contribution by private companies is not affected by the patent granting. These results confirm Sampat and Williams (2019) intuition about potential differences in licensing efficiency between non-profit and commercial firms. In particular, while the sudden introduction of "fencing" against scientific knowledge does not affect its use by commercial firms, it is detrimental to research conducted in universities and public research centres.

[11]The models are estimated using both Ordinary Least Square (OLS) estimation with multiple level of fixed effect (Columns 1 and 3) and Poisson Quasi Maximum Likelihood (PQML) estimation (Columns 3 and 4)

Table 4: Effect of patenting on follow-up scientific research (**RQ1**)

DEPENDENT VARIABLES:	CIT_PUBLIC		$CIT\_PRIVATE$	
	(OLS)	(PQML)	(OLS)	(PQML)
	(1)	(2)	(3)	(4)
$PATENT\_IN\_FORCE$	-2.269***	-0.239***	-0.006	-0.117
	(0.660)	(0.058)	(0.075)	(0.121)
PAT_WINDOW	-1.228*	-0.159***	-0.121	-0.290**
	(0.619)	(0.034)	(0.070)	(0.102)
$PAPER\_AGE$		0.460***		0.377***
		(0.027)		(0.048)
$PAPER\_AGE^2$		-0.051***		-0.060***
		(0.004)		(0.008)
$PAPER\_AGE^3$		0.001***		0.002***
		(0.000)		(0.000)
$Constant$	10.960***		0.550***	
	(0.394)		(0.045)	
Observations	4439	4375	4439	3266
R^2	0.869		0.644	
ll		-9484.1		-2206.5

Note: Column 1 and 3 are estimated using a linear model with multiple levels of fixed effects. Column 2 and 4 are estimated using the Poisson Quasi-Maximum Likelihood model. Standard errors are reported in parenthesis.
Legend:* p<0.05, ** p<0.01, *** p<0.001

Table 5: Additional effect of knowledge applicability (**RQ2**)

DEPENDENT VARIABLES:	*CIT_PUBLIC* (OLS) (1)	*CIT_PRIVATE* (OLS) (2)
PATENT_IN_FORCE	-3.175***	-0.183*
	(0.775)	(0.0877)
PATENT_WINDOW	-1.248	-0.154*
	(0.660)	(0.0747)
*PATENT_IN_FORCE*APPLIED_KNOWLEDGE*	1.519*	0.311***
	(0.634)	(0.0718)
Constant	11.46***	0.596***
	(0.421)	(0.0476)
Number of observations	4160	4160
R^2	0.871	0.649

Note: Estimations are performed using a linear model with multiple levels of fixed effects. Standard errors are reported in parenthesis.
Legend:* p<0.05, ** p<0.01, *** p<0.001

Table 5 reports the estimated results related to our second research question (**RQ2**). The coefficient of the interaction term reported in column 1 of table 5 is positive and significant. This result indicates that the negative effect of the patenting can be mitigated if the scientific knowledge of a publication is more applied. Since the coefficient of the interaction reported in column 2 of table 5, this result is confirmed for subsequent scientific research developed by private firms.

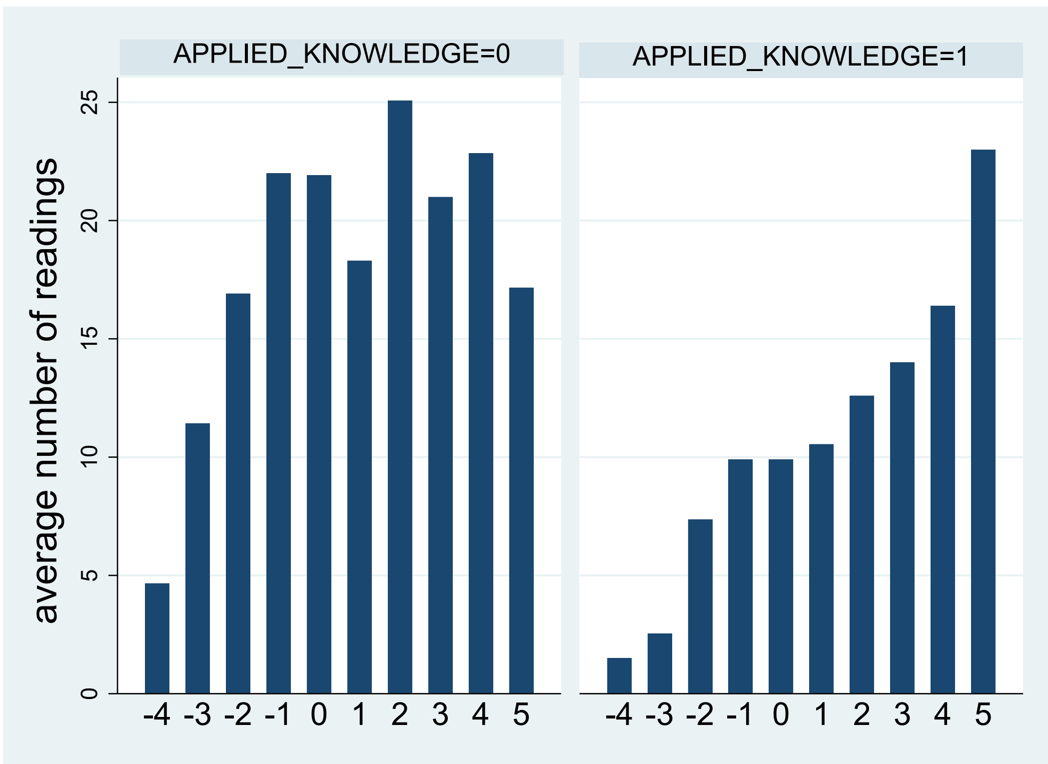
From a public policy point of view, one of the justifications for the patent system is that patents encourage disclosure, and more generally, generate rapid and wide diffusion of technical information on the most recent inventions (Machlup, 1958). In line with this theory, Mazzoleni and Nelson (1998) suggest that universities may also publicize potentially commercial research through their patents. In the case of PPP, patent visibility might spill over to publication generating an "advertising effect". In our setting, more applied publications might benefit from this "advertising effect" of patents and result in an increased number of citations. To test this, we use Scopus data provided by PlumX Metrics[12] to retrieve the number of annual accesses to papers to proxy for the evolution of publication visibility over time. Data availability constrains this check to the sub-sample of publications paired to a patent granted in 2011 or 2012[13].

Figure 11 shows the average number of reads received by applied and non-applied scientific publication in the years before and after the patent grant, and zero is the moment of the patent award. The figure shows that the visibility of non-applied papers does not change significantly after the granting of a patent; whereas, the reverse is true for applied publications. Figure 11 confirms Mazzoleni and Nelson's (1998) intuition about the role played by patents as advertising mechanisms for the paired publication. Even more interesting is the evident difference in visibility between the two groups even before the patent granting. This difference suggests a "freedom to advertise effect", rather than a simple "advertising effect". Although the scientific publication should play an essential role in disclosure, it would seem that the difference in visibility across the two groups could be driven by the tendency to hide pieces of applied knowledge before the patent is awarded. The need for hiding vanishes as knowledge becomes protected by a formal IPR and the

[12]See `https://blog.scopus.com/topics/plumx-metrics`

[13]This corresponds to 81 publications (25% of the sample)Data availability constraints is due to missing values for older periods.

Figure 11: Visibility differences



inventor becomes free to advertise his invention with no-risk of copying.

5 Conclusions

The emergence and consolidation of a "pro-market" policy regime increasingly focused on the rapid commercialisation of scientific findings has prompted a lively debate on its possible unintended effects on the research agenda of universities and other public research organizations, as well on its other potential benefits and drawbacks. So far, most of the literature has been focusing on the direct effects of this new regime of science commercialization on social welfare (Bornmann, 2013; Shane, 2004b). In fact, consumers can benefit of scientific advantages only in case these are converted into a practical application. Nevertheless, the key-role of basic long term *blue sky* research for the entire innovation process is corroborated by a large amount of quantitative and qualitative evidence (Bush, 1945; Nelson, 1992). This raises the concern that strong policies for encouraging industrial exploitation of basic research may lead to shift research priorities of universities and other public organizations from long term research towards short term applications. In addition, scientific public knowledge has historically been a *not-rival* good since after its public disclosure no further restrictions on its use can be imposed. This lead to concerns that reduced access today can impact on the rate of follow-up scientific research.

This report focuses on the consequences of basic science commercialisation on basic science itself. While most of the literature on university-industry relations studies several mechanisms that make science more easily exploitable by the market; we focus on the role of patenting of scientific discovery. Traditionally science and technology have been characterised by different institutions and norms. As regards science, these institutions and norms have been promoting the open science paradigm (David, 2004; Merton, 1973; Partha and David, 1994). The pressure to increase science commercialisation has promoted the "privatisation" of scientific knowledge using patents and changed the balance towards the openness of science. In this report, we focus on the role of patenting on top scientists publishing behaviour and on the development of subsequent scientific knowledge. We chose to focus the analysis on top scientists because of the existing evidence on top scientists being very active in producing academic knowledge and commercialising their research output (Zucker and Darby, 1997, 2001). Furthermore, we gave particular attention to the field of bio-science as the practice of patenting basic knowledge is widely spread in this area (Holman and Munzer, 2000).

The report consists of three sections reporting different pieces of evidence.

In section 2 we analysed the trend characterising the patent activity of top scientists. Our results show that half of top performers scholars are also patent inventors. The attitude to make proprietary, via patenting, the output of basic research has drastically increased starting from 1980s'. Besides, we provide evidence that the share of each assignee type over time shows a substantial increase of universities as patent assignees, consistently with the increasing incentives for universities to directly engage in the commercialisation of scientific research.

In section 3 we restricted the analysis within the bio-science field, and we investigated whether different dynamics between *pure scientists* and *scientists-inventors* publishing activity emerge. Also, we explore the possibility that some changes occur after the patent grant. We provided descriptive evidence that *scientists-inventors* publish relative more than *pure scientists*, but no evidence of better quality research emerges. One important difference is in the inclination of top scientists to freely disclose research output. While *pure scientists* are on average more willing to give an open-access status to scientific production, *scientists-inventors* become more flexible on accessibility conditions after the patent award. This result suggests that scholars tend to freely disclose their scientific production only when a formal IPR protects them.

In section 4 we focus on a specific case of bio-science, namely *research on cancer detection* to provide econometric evidence on the effect of patenting on subsequent scientific development. Using PPP, as firstly proposed by Murray (2002), we showed that granting patents on pieces of knowledge previously publicly available has a significant negative effect on the rate of follow-up scientific research. This effect is mitigated (i.e. less negative) when the protected knowledge is more applied. Also, this result seems compatible with the previously mentioned tendency to "hide" more research output closer to market before patent grant. In this sense, the granted patent is a tool that gives scientists a *freedom to advertise*.

All in all, the results provided by this report should stimulate policy makers to consider a possible unintended and long-term detrimental effect of the current policy regimes favoring science commercialisation on social welfare alongside the positive direct effect usually evoked within the literature. In particular, our evidence suggest that there is indeed the serious risk that these policies will ultimately result in a set of incentives that will determine a major shift in the research agenda of top scientists from basic to more applied research. If this is the case, a less sanguine and more sobering approach in the set of public policies that affect the commercialisation of research results of universities is probably in order.

References

Arora, A., S. Belenzon, and A. Patacconi (2015): "Killing the golden goose? The decline of science in corporate R&D," Tech. rep., National Bureau of Economic Research.

Audretsch, D. and D. Göktepe-Hultén (2015): "University patenting in Europe," *The Chicago handbook of university technology transfer and academic entrepreneurship*, 188–217.

Balconi, M., S. Brusoni, and L. Orsenigo (2010): "In defence of the linear model: An essay," *Research Policy*, 39, 1 – 13.

Bok, D. (2009): *Universities in the marketplace: The commercialization of higher education*, vol. 49, Princeton University Press.

Bornmann, L. (2013): "What is societal impact of research and how can it be assessed? A literature survey," *Journal of the American Society for information science and technology*, 64, 217–233.

Bush, V. (1945): "Science: The endless frontier," *Transactions of the Kansas Academy of Science (1903)*, 231–264.

Callaert, J., B. Van Looy, A. Verbeek, K. Debackere, and B. Thijs (2006): "Traces of Prior Art: An analysis of non-patent references found in patent documents," *Scientometrics*, 69, 3 – 20.

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

Clarysse, B., V. Tartari, and A. Salter (2011): "The impact of entrepreneurial capacity, experience and organizational support on academic entrepreneurship," *Research policy*, 40, 1084–1093.

Cockburn, I. M. and R. M. Henderson (1998): "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery," *The Journal of Industrial Economics*, 46, 157–182.

Cohen, W. M. and D. A. Levinthal (1989): "Innovation and Learning: The Two Faces of R D," *The Economic Journal*, 99, 569–596.

David, P. A. (2004): "Can 'Open Science'be protected from the evolving regime of IPR protections," *Journal of Theoretical and Institutional Economics*, 160, 1–26.

David, P. A. et al. (2000): "The digital technology boomerang: New intellectual property rights threaten global 'open science'," in *World Bank Conference Paper*.

Deci, E. L., R. Koestner, and R. M. Ryan (2001): "Extrinsic rewards and intrinsic motivation in education: Reconsidered once again," *Review of educational research*, 71, 1–27.

Dent, C., P. Jensen, S. Waller, and B. Webster (2006): "Research Use of Patented Knowledge," OECD Science, Technology and Industry Working Papers No. 2006/02, OECD Publishing, Paris.

Dreyfuss, R. (2004): "Protecting the public domain of science: Has the time for an experimental use defense arrived?" *Arizona Law Review*, 46, 457–472.

Etzkowitz, H. and L. Leydesdorff (2000): "The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations," *Research policy*, 29, 109–123.

Fleming, L. and O. Sorenson (2004): "Science as a map in technological search," *Strategic Management Journal*, 25, 909–928.

Fuller, A. W. and F. T. Rothaermel (2012): "When stars shine: The effects of faculty founders on new technology ventures," *Strategic Entrepreneurship Journal*, 6, 220–235.

Gambardella, A. et al. (1995): *Science and innovation: The US pharmaceutical industry during the 1980s*, Cambridge University Press.

Garfield, E. (1979): "Is citation analysis a legitimate evaluation tool?" *Scientometrics*, 1, 359–375.

Grimpe, C. and H. Fier (2010): "Informal university technology transfer: a comparison between the United States and Germany," *The Journal of Technology Transfer*, 35, 637–650.

Hamilton, K. S. (2003): "Subfield and level classification of journals," *CHI Research*.

Heller, M. A. and R. S. Eisenberg (1998): "Can patents deter innovation? The anticommons in biomedical research," *Science*, 280, 698–701.

Holman, M. A. and S. R. Munzer (2000): "Intellectual Property Rights in Genes and Gene Fragments: A Registration Solution for Expressed Sequence Tags," *Iowa Law Review*, 85, 735–750.

Huang, K. G. and F. E. Murray (2009): "Does patent strategy shape the long-run supply of public knowledge? Evidence from human genetics," *Academy of Management Journal*, 52, 1193–1221.

Kline, S. J. (1985): "Innovation is not a linear process," *Research management*, 28, 36–45.

Machlup, F. (1958): *An economic review of the patent system*, 15, US Government Printing Office.

Mazzoleni, R. and R. R. Nelson (1998): "The benefits and costs of strong patent protection: a contribution to the current debate," *Research policy*, 27, 273–284.

Merton, R. K. (1973): *The Sociology of Science*, University of Chicago Press.

Mowery, D. C., R. R. Nelson, B. N. Sampat, and A. A. Ziedonis (2001): "The growth of patenting and licensing by US universities: an assessment of the effects of the Bayh–Dole act of 1980," *Research policy*, 30, 99–119.

Murdock, K. (2002): "Intrinsic motivation and optimal incentive contracts," *RAND Journal of Economics*, 650–671.

Murray, F. (2002): "Innovation as co-evolution of scientific and technological networks: exploring tissue engineering," *Research Policy*, 31, 1389–1403.

Murray, F. and S. Stern (2007): "Do formal intellectual property rights hinder the free flow of scientific knowledge?: An empirical test of the anti-commons hypothesis," *Journal of Economic Behavior & Organization*, 63, 648–687.

Narin, F., K. S. Hamilton, and D. Olivastro (1997): "The increasing linkage between US technology and public science," *Research Policy*, 26, 317–330.

Nelson, R. R. (1992): *Technology and the Wealth of Nations*, Stanford University Press, chap. What is "Commercial" and What is " Public" About Technology, and What Should Be?

Partha, D. and P. A. David (1994): "Toward a new economics of science," *Research Policy*, 23, 487 – 521, special Issue in Honor of Nathan Rosenberg.

Posner, R. A. (2000): "An economic analysis of the use of citations in the law," *American Law and Economics Review*, 2, 381–406.

Price, D. J. d. S. (1965): "The scientific foundations of science policy," .

Rosenberg, N. (1990): "Why do firms do basic research (with their own money)?" *Research Policy*, 19, 165 – 174.

Sampat, B. and H. L. Williams (2019): "How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome," *American Economic Review*, 109, 203–36.

Shane, S. (2004a): "Encouraging university entrepreneurship? The effect of the Bayh-Dole Act on university patenting in the United States," *Journal of Business Venturing*, 19, 127–151.

Shane, S. A. (2004b): *Academic entrepreneurship: University spinoffs and wealth creation*, Edward Elgar Publishing.

Siegel, D. S., R. Veugelers, and M. Wright (2007): "Technology transfer offices and commercialization of university intellectual property: performance and policy implications," *Oxford review of economic policy*, 23, 640–660.

Siegel, D. S., D. A. Waldman, L. E. Atwater, and A. N. Link (2003): "Commercial knowledge transfers from universities to firms: improving the effectiveness of university–industry collaboration," *The Journal of High Technology Management Research*, 14, 111–133.

Sorenson, O. and T. E. Stuart (2001): "Syndication networks and the spatial distribution of venture capital investments," *American journal of sociology*, 106, 1546–1588.

Squicciarini, M., H. Dernis, C. Criscuolo, et al. (2013): "Measuring patent quality: Indicators of technological and economic value," Tech. rep., OECD Publishing.

Stern, S. (2004): "Do scientists pay to be scientists?" *Management science*, 50, 835–853.

Torero, M., M. R. Darby, and L. G. Zucker (2001): "The importance of intellectual human capital in the birth of the semiconductor industry," *Management Science*, 48.

Zucker, L. G. and M. R. Darby (1997): "Individual action and the demand for institutions: Star scientists and institutional transformation," *American Behavioral Scientist*, 40, 502–513.

——— (1998): "Entrepreneurs, star scientists, and biotechnology," *NBER Reporter Online*, 7–10.

——— (2001): "Capturing technological opportunity via Japan's star scientists: Evidence from Japanese firms' biotech patents and products," *The journal of Technology transfer*, 26, 37–58.

A Appendices

A.1 Empirical strategy

To answer our research questions, we developed an econometric model to isolate the effect of being granted a patent on the rate of follow-on public knowledge. Following the literature (Price, 1965; Posner, 2000), we use number of yearly citations received by a (paired) publication to measure follow-on public knowledge. To understand the rationale behind our identification strategy, we would highlight two points. First, the so-called grant lag identifies the time elapsed between date of filing the patent application and date of patent award. Since there is correspondence between patent application and paper publication, the grant lag identifies the period when the new knowledge is freely available and usable. After a patent has been granted, it becomes risky (in terms of infringement) for scientists to continue to exploit the knowledge without obtaining a license. Second, the average grant lag in our sample is 4.5 years with considerable heterogeneity ($\sigma = 755.23$). Figure 12 is a graphical representation of the grant lag distribution in our sample.

The high variation in grant lags allows us to consider award of a patent as an exogenous shock to the corresponding publication during "its life". In this setting we can implement a difference-in-difference model where the annual citations received by papers associated to patents with longer grant-lags constitute the "control" for citations to a paper associated to a patent with a shorter patent lag. Figure 13 provides a graphical representation of the identification strategy.

Figure 12: Distribution of grant lag

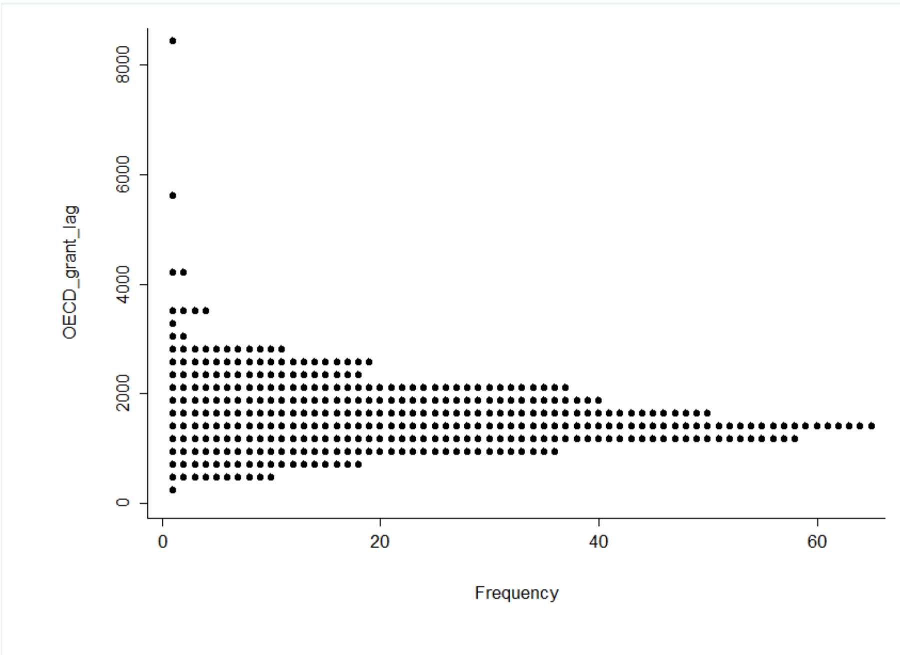
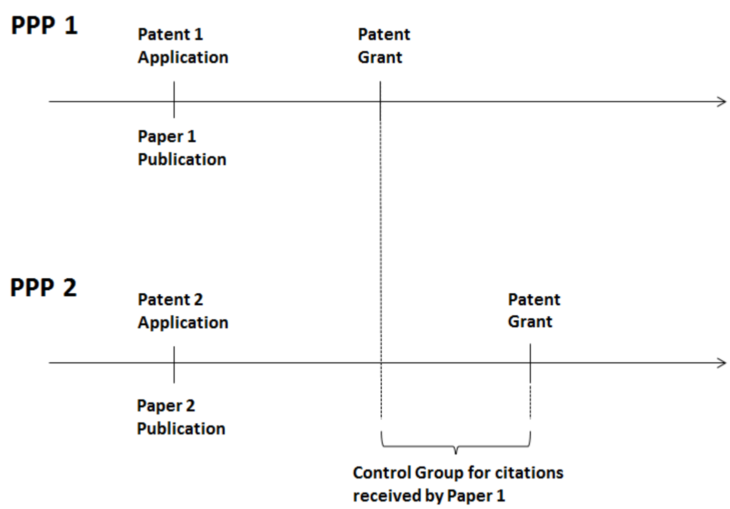


Figure 13: Scheme of the patent-paper pair identification strategy



A.2 Descriptive statistics

Tables A.1 and A.2 present the summary statistics and the correlations, respectively.

Table A.1: Summary statistics

	count	mean	sd	min	max
CIT_PUBLIC	4455	9.657	24.22	0	306
CIT_PRIVATE	4455	0.535	1.666	0	28
PAT_WINDOW	4455	0.0837	0.277	0	1
PATENT_IN_FORCE	4455	0.513	0.500	0	1
PAPER_AGE	4455	6.170	4.493	0	29
APPLIED_KNOWLEDGE	4176	0.459	0.498	0	1
Observations	4455				

Table A.2: Cross-correlation table

Variables	*CIT_PUBLIC*	*CIT_PRIVATE*	*PAT_WINDOW*	*PATENT_IN_FORCE*	*PAPER_AGE*	*APPLIED_KNOWLEDGE*
CIT_PUBLIC	1.000					
CIT_PRIVATE	0.763	1.000				
PAT_WINDOW	0.023	-0.002	1.000			
PATENT_IN_FORCE	-0.022	-0.084	-0.310	1.000		
PAPER_AGE	-0.024	-0.090	-0.091	0.712	1.000	
APPLIED_KNOWLEDGE	-0.043	-0.050	0.024	-0.055	-0.102	1.000