



RESEARCH ARTICLE

# Technological impact of funded research: A case study of nonpatent references

Justin Quemener<sup>1,2</sup> , Luis Miotti<sup>1</sup>, and Abdelghani Maddi<sup>1,3</sup> 

<sup>1</sup>Observatoire des Sciences et Techniques, Hcéres, Paris, France

<sup>2</sup>Université Côte d'Azur, Nice, France

<sup>3</sup>GEMASS – CNRS – Sorbonne University, Paris, France

an open access  journal



**Keywords:** citation impact, knowledge flow, nonpatent literature, nonpatent references, research funding, technological impact

## ABSTRACT

Research funding is essential to expand knowledge, foster innovation, and address the complex challenges that shape our future. The scientific literature has extensively addressed the relationship between research funding and academic impact. More recently, several studies have analyzed the technological impact of funded research as measured through citations in patents, known as *nonpatent references* (NPRs). But there remains much to know about NPRs and the multiplication of case studies is necessary to characterize them. Here we analyze a sample of 7,065 publications funded by the French Foundation for Medical Research (FRM) and the citations of these publications in patents. This study shows the high scientific and technological impacts of FRM funding. Indeed, the publications funded by FRM that are cited in patents are 3.5 times more frequently cited by other publications than the global average (for funded publications in the entire database, this ratio is 2.6). Furthermore, our results also indicate that USPTO patents citing these publications exhibit high-quality indicators. Moreover, five of these patents have led to approved drug products by the United States Food and Drug Administration (FDA). This study provides further evidence of the positive influence that research funding can have on both scientific and technological advancements.

## 1. INTRODUCTION

The impact of public research funding on innovation is generally measured by identifying the public institutions among patent applicants. Since the 1980s in the United States and the 1990s in Europe, several policies have encouraged the patenting by academic institutions of their inventions. These policies and the dynamics of innovation in a number of sectors have significantly increased the number of patents owned by public institutions (Henderson, Jaffe, & Trajtenberg, 1998; Mowery & Sampat, 2004). Nevertheless, the impact of academic research on innovation can be much broader, in particular by contributing to firms' inventions (Wang & Li, 2021). Research remains a risky activity, and private firms still rely a lot on public research (Cross, Rho et al., 2021; Hausman, 2022). In addition, the academic sector represents the vast majority of publications, but only a minority of patents, in the world.

Measuring the impact of scientific research on innovation can be a complex challenge. Various approaches can be employed to assess this impact, such as the employment of young doctors in the private sector, the mobility of public researchers to the private sector, or the

Citation: Quemener, J., Miotti, L., & Maddi, A. (2024). Technological impact of funded research: A case study of nonpatent references. *Quantitative Science Studies*, 5(1), 170–186. [https://doi.org/10.1162/qss\\_a\\_00281](https://doi.org/10.1162/qss_a_00281)

DOI:  
[https://doi.org/10.1162/qss\\_a\\_00281](https://doi.org/10.1162/qss_a_00281)

Peer Review:  
[https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss\\_a\\_00281](https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00281)

Supporting Information:  
[https://doi.org/10.1162/qss\\_a\\_00281](https://doi.org/10.1162/qss_a_00281)

Received: 16 April 2022  
Accepted: 21 November 2023

Corresponding Author:  
Abdelghani Maddi  
[abdelghani.maddi@cnrs.fr](mailto:abdelghani.maddi@cnrs.fr)

Handling Editor:  
Vincent Larivière

Copyright: © 2024 Justin Quemener, Luis Miotti, and Abdelghani Maddi. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.



creation of start-ups from the public sector. Nevertheless, these approaches are mostly based on survey data, and can be used mainly for case studies. Because they do not have a normalized database, these approaches do not generally allow building comparable indicators at the international level or across technologies. Another approach appears to be promising: the use of nonpatent references (NPRs) citations in patents. Based on patent data, this approach can potentially help to build comparable indicators across technologies or countries, for instance. NPRs refer to the scientific or technical literature that is cited within patents. NPRs can function as a valuable proxy for assessing the extent to which scientific literature is employed in patents, as they offer insights into the scientific knowledge and literature that hold relevance and may in some cases impact the development of the claimed innovations. Thus, NPRs can offer information about the scientific foundation and the state of the art related to the patented inventions. Furthermore, the quality and credibility of the cited publications can reflect the robustness and reliability of the scientific basis supporting the patented inventions.

When evaluating the outputs of an institution, whether from the public or the private sector, it is usual to measure the production and quality of scientific publications and those of patents originating from this institution separately, without analyzing the broader influence of institutions' research on technology, including on patents granted by other institutions. By analyzing an institution's NPRs (the institution's scientific publications cited in patents), in addition to more traditional indicators, such as scientific publications and patents, one can gain insights into the extent to which the institution's research is influencing and contributing to innovation. Considering both publications and NPRs allows for a comprehensive assessment of an institution's research output and its influence on the broader scientific and technological landscape. It provides a more holistic view of how an institution's research contributes to innovation, leveraging the information contained within both the patents and the cited scientific literature.

In a recent literature review, Velayos-Ortega and López-Carreño (2023) showed a significant increase in the analysis of NPRs in recent years, particularly starting from 2020. This rise can be attributed to remarkable advancements that were made in identifying and matching scientific publications cited in patents with bibliometric databases. The database developed by Marx and Fuegi (2020, 2022) emerged as a highly suitable resource for conducting such analyses, allowing for the differentiation between front page and in-text scientific citations. This capability enhanced the precision and granularity of NPR analysis, providing a more comprehensive understanding of the specific references made within patents and their corresponding scientific publications. However, the work of Velayos-Ortega and López-Carreño (2023) has highlighted that there remains a gap in our understanding regarding the characteristics of scientific publications that are cited in patents. Despite the increase in the analysis of NPRs, further research is needed to explore and comprehend the specific attributes and qualities of these cited scientific publications. This gap highlights the ongoing need for comprehensive studies that delve into the nature and impact of scientific literature in the patenting process.

This article aims to address the following research question: What are the specific characteristics in terms of citation impact of publications cited in patents? Conversely, do patents citing academic literature exhibit higher quality? By analyzing the scientific impact of publications and considering the role of funding, we seek to gain a deeper understanding of the incorporation of scientific knowledge into technology and its potential positive effects on citation. Furthermore, this study also examines the influence of science on patent quality.

To accomplish this, we proposed a novel approach for calculating the traditional mean normalized citation score (MNCS) indicator, taking into account whether the article received funding and/or was cited in patents. Regarding patents, it uses four different indicators of quality (family size, claims, originality, and renewal) to have a broader view of the impact of

science on quality. It also compares the characteristics of patents citing science at the United States Patent and Trademark Office (USPTO) and at the European Patent Office (EPO), the latter being much less frequently analyzed in the literature. Finally, it also links citing patents with drugs (using the Orange Book database), which has not often been studied in the literature (to our knowledge, only Du, Li et al. (2019) have studied the whole knowledge transfer spectrum from scientific articles to drugs). These methodologies allowed us to assess patents that cite scientific publications in a more robust and comprehensive manner, introducing new insights that enhance the evaluation of the relationship between patents and scientific literature.

It is important to note that this study does not aim to directly assess the impact of NPRs on inventive activity and technological development. Instead, our approach is divided into two distinct phases. First, we seek to characterize the publications receiving citations from patents. Second, we delve into characterizing the patents that cite these NPRs to better understand how scientific knowledge is integrated into the patenting process. To fully analyze the impact of NPRs on technology, a more in-depth investigation would be required, including examining the nature of the citations, their context, their applicability, and how all these factors may eventually translate into a tangible contribution to patented innovation.

This study is part of a larger research project supported by the Foundation for Medical Research (FRM; <https://www.frm.org/en>), a charitable organization established in 1947 to fund medical research in France. Thus, this analysis is carried out both for the FRM and for the entire database of publications used (Web of Science [WoS]). To ensure comparability, we also created a control group composed of publications similar to those funded by the FRM. On the patent side, we conducted the same exercise by comparing the quality of patents citing FRM-funded publications to a control group, using the USPTO and EPO databases.

We organize the rest of the article as follows: an overview of the existing literature on the issue, presentation of the data and method, the main results, and finally a conclusion and discussion.

## 2. LITERATURE REVIEW ON NPRs

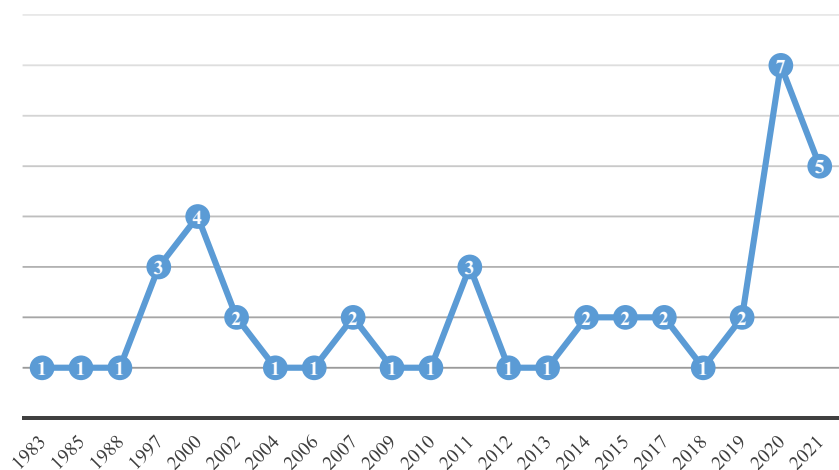
The literature has extensively explored the relationship between scientific research funding and academic impact, with the consensus being that research receiving funding tends to exhibit greater academic impact and produces high-quality publications (Álvarez-Bornstein & Bordons, 2021; Jowkar, Didegah, & Gazni, 2011; Leydesdorff, Bornmann, & Wagner, 2019; Roshani, Bagheryllooieh et al., 2021; Yan, Wu, & Song, 2018). Studies examining the impact of funding on technology have revealed noteworthy findings. Some research suggests that the impact of funding is more pronounced in the realm of human resource capability development when compared to research and technological output (Garg, Gupta et al., 2005). In fact, receiving project funding has been associated with heightened technological impact (Sohn, Gyu Joo, & Kyu Han, 2007). Furthermore, in the context of innovation financing, a study demonstrated that early-stage awards roughly double the probability of a firm securing subsequent venture capital investments (Howell, 2016). Conversely, another study highlighted the presence of financing constraints as significant barriers to innovation, underscoring the critical role of funding in fostering innovative endeavors (Silva & Carreira, 2012). These collective insights underscore the multifaceted influence of funding on technology, innovation, and organizational development. In this literature review, our primary focus centers on NPRs and their analysis within the existing body of literature.

Several studies have used the front-page NPRs in patent documents, which include scientific publications, to identify the knowledge sources the inventions build upon. The first study

to examine scientific citations in patents was conducted by Carpenter and Narin (1983), who analyzed the citations made by patent examiners to journals, reports, and books. Two years later, Narin and Noma (1985) analyzed the citations of “non-patent resources” in patents and found that they represented 0.3 references per patent. Since then, several dozen papers have been published on the topic. Figure 1, based on Velayos-Ortega and López-Carreño’s (2023) review, illustrates the number of studies conducted over the years that have analyzed scientific citations in patents. A sharp increase is observed in the late 1990s, which picks up again from 2020 with seven recent publications on the topic. Next is a brief literature review of the main studies that are directly related to our own study.

Roach and Cohen (2013) found that NPRs are a better measure of knowledge originating from public research than patent references. Gazni and Ghaseminik (2019) showed that the top 1% highly cited patents in the USPTO relied increasingly on science during 1992–2016. They also showed that top patents with inventors from countries such as the United States, the United Kingdom, and France seem to rely more on science than applicants from other countries (although differences in technological structures should be taken into account). Ahmadpoor and Jones (2017) measured the distance between science and technology for several fields. They showed that health-related scientific fields, such as “Virology,” “Chemistry, Clinic and Medicinal,” “Cell Biology,” “Biotechnology,” and “Immunology,” are close to the technological frontier, which means that the publications in these topics are often cited in USPTO patents.

To measure knowledge transfer from science to innovation more effectively, Bryan, Ozcan, and Sampat (2020) advocated using in-text citations instead of front-page citations. They showed, for USPTO patents, that in-text citations are closer to the concept of invention and that citations from the front page are mainly added by applicants to comply with the legal “duty of disclosure.” They also showed that only 31% of in-text citations appear on the front page, thereby confirming that these two types of citations are different. Marx and Fuegi (2020) completed these findings by showing that in-text citations contain fewer self-citations, and that applicants from the nonprofit sector have a higher propensity to add in-text citations, whereas firms have a higher propensity to add front-page citations. Finally, Wang and Verberne (2021) showed, for a set of USPTO biotech patents, that in-text references are on average more basic, receive more scientific citations, are less interdisciplinary, and are slightly more novel than front-page references.



**Figure 1.** Number of studies that have analyzed NPRs by year (based on results from Velayos-Ortega and López-Carreño (2023)).

Ahmadpoor and Jones (2017) showed that only 5% of the scientific publications from WoS are cited at least once on the first page of USPTO patents. Their study also showed the positive correlation between the number of forward scientific citations received by publications and their propensity to be cited in patents, which is in line with the results of most studies (Bikard & Marx, 2020; Poege, Harhoff et al., 2019; Popp, 2017). Veugelers and Wang (2019) and Ke (2020) showed the positive correlation between the novelty of publications and their propensity to be cited in patents. Ke (2020) also found a positive correlation between the basicness of biomedical publications and their propensity to be cited in patents, and Wang and Verberne (2021) found an inverted U-shape effect of basicness on patent forward citations.

Assessment studies of public funding of science have also used nonpatent front-page references recently to measure the impacts of public investments on innovation. In the field of biomedical research, Li, Azoulay, and Sampat (2017) showed that 8% of NIH grants generate a patent directly, and 31% of grants generate an article that is subsequently cited by firms' patents. This result underscores the interest of taking indirect effects of the funding into account. Fleming, Greene et al. (2019) also showed that the impact of government-funded research should be measured taking indirect effects into account, through patent and publication citations in patent documents. From 1926 to 2017, some 70,000 patents were assigned to the U.S. government, and patents that relied directly or indirectly on government support represented around 1,135,000 patents, along with 100,000 patents that could be linked to publications acknowledging government support. Moreover, Popp (2017) showed that, in the clean-energy research field, government institutions play a significant role, as their publications are more likely to be cited in energy patents than those of any other institutions (including companies and universities).

Some studies have gone further and tried to measure economic impacts. Du et al. (2019) measured linkages in the whole knowledge transfer spectrum in the drug sector (drugs-patent-papers-grants). They showed that 75% of scientific papers cited by drug patents in the U.S. Food and Drug Administration (FDA) Orange Book database come from the public sector, and that 90% of them have received government support. They also showed that the average number of papers cited in drug patents has increased from five in the 1990s to 30 in the 2010s, and that the lag between the publication and the grant date of the patent (called here the "maturity of the scientific knowledge") doubled between 1993 and 2015. Finally, Chen, Kan, and Tung (2016) showed, for a set of Taiwanese electronic firms, that the citation of scientific publications by a patent has a strong and positive effect on the productivity of the patent-owning firm.

### 3. MATERIAL AND METHODS

#### 3.1. Patent Data

The patent data were extracted from the OST (Observatoire des Sciences et Techniques) patent database, built from PATSTAT and enriched by the OST. The PATSTAT database was created by the EPO with support from the OECD in particular. The EPO updates and distributes the entire database twice a year (April and October). The extracted information is based on the version of PATSTAT from April 2019 and takes account of all requests published up to February 2019.

PATSTAT contains the records of patent filings after publication of the application (i.e., 18 months after the date of the first filing). Only published requests therefore appear in the indicators. PATSTAT covers 80 national and regional patent offices around the world.

In this study, we focused on patents from the EPO and USPTO.

### 3.2. Scientific Publications Data

The publications data were extracted from the OST in-house database. They include five indexes of WoS available from Clarivate Analytics (Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (AHCI), Conference Proceedings Citation Index (CPCI-SSH), and Conference Proceedings Citation Index (CPCI-S)), and corresponds to WoS content indexed through the end of March 2019. Indicators were calculated for four types of publications (articles, letters, reviews, and conference proceedings) and only for Natural Sciences and Engineering (NSE) WoS subject categories (178 subject categories out of 254)<sup>1</sup>. The choice of working on NSE only was motivated by the fact that NPRs are quite rare in the Social Sciences and Humanities.

The FRM data set is made up of lists of publications identified by the FRM and provided to the OST for the purposes of a dedicated study. This therefore ensures exhaustive identification of publications that have received funding from this agency. The data set contains 8,608 publications, including 7,065 with a DOI.

### 3.3. Linking Patents to Publications

Linking patents to publications is not an easy task because NPRs do not have a defined format, unlike patent references; they are often incomplete and contain typo-errors. First, it is necessary to clean the data using a matching algorithm. Although it is quite easy to clean data for a subset of patents in a particular field, it is more difficult to find a given subset of publications in patent citations because in this case, all the patent data have to be cleaned for the different offices of interest. Extracting in-text patent references is even more complex, because the references are not contained in a specific row, as they are for front-page citations. For this, it is necessary to rely on huge computer resources. This is the reason why we decided to use the data kindly provided by Marx and Fuegi (2020), which contains front-page NPR data for several offices and in-text citations for the USPTO. DOIs of publications were used to match NPRs with the WoS database. Given the gaps related to the completeness of the DOIs in the WoS database before 2009, we only worked on the recent period 2009–2018.

To build our database, we tagged all the publications in the WoS database, from 2009 to 2018, as “NPR” and “non-NPR” by matching with the Marx and Fuegi (2020) patent database. As patent filing is rare in the Social Sciences and Humanities, only the NSE WoS subject categories were tagged (with 11,341,437 publications over the period). Then we calculated the normalized citation scores for NPR and non-NPR publications funded by the FRM (316 NPRs out of 7,065 publications), for all the NPRs in the WoS database (315,164 publications), and for non-NPR ones (11,026,291 publications), taking the funding information into account using the “acknowledgements” field provided in the WoS database.

In this stage, we considered that if a citation to a scientific publication from either an EPO application or an USPTO patent did not have a different interpretation, then we merged the citations from both offices. It is importance to notice that for all our analysis, we used patent applications from the EPO (to have a more significant sample for the patent analysis), and granted patents for the USPTO (because scientific citations only appear in the granted patent publications).

---

<sup>1</sup> See [https://images.webofknowledge.com/images/help/WOS/hp\\_research\\_areas\\_easca.html](https://images.webofknowledge.com/images/help/WOS/hp_research_areas_easca.html).



### 3.4. Linking Patents to Drugs from the Orange Book Database

The publication *Approved drug products with therapeutic equivalence evaluations* (commonly known as the *Orange Book*<sup>2</sup>), identifies drug products approved on the basis of safety and effectiveness by the FDA under the Federal Food, Drug, and Cosmetic Act (the FD&C Act) and related patent and exclusivity information.

We used the Orange Book data files from December 2020. The “patent.txt” file provided the patent publication number of the approved products that we could match with our data set.

### 3.5. Control Sample of Citing Patents

To determine the quality of USPTO granted patents and EPO patent applications citing the publications funded by the agency, we created a control sample of both USPTO and EPO patents. In this stage, we analyzed separately the two sets of citing patents from the USPTO and the EPO, with their respective control sample, because the values of quality indicators are on average significantly different in each office and are thus not directly comparable. For example, the average value of claims, where the average value of renewal is higher in EPO; see annex D1 in the Supplementary material for a comparison across offices of the four quality indicators used.

Firstly, all patents filed from 2004 to 2018, with the same combination of IPC7 technology classes as at least one of the patents in the funding agency data set, were included in the sample.

The difficulty is that even after controlling the combination of technologies in the first step, the control sample might still display characteristics that differ from the treated sample (the patents citing FRM-funded articles), preventing the correct measurement of the treatment effect. That is why we used the propensity score matching method to select for each observation of both treated samples (USPTO and EPO) the closest observation in their respective control samples, based on a set of covariates (Corsini & Pezzoni, 2023; Dehejia & Wahba, 2002; Rubin, 2001).

The propensity scores were estimated using a logistic regression. The equation of this logistic regression is available in the Supplementary material. The selected three covariates are the filing year, the technology class IPC4, and the institutional sector of the applicant, all used as dummies. We selected these variables because our data exploration showed that they have a significant impact on the values of the quality indicators. For the filing year, the effect is mechanical, especially for renewal and to a lesser extent on the size of the family for the last year considered. The technology class is the variable whose impact is often the most significant on quality indicators. Finally, the institutional sector also plays a role, in particular in the strategies for extending applications abroad. Because some observations may not have a known value for one of the quality indicators, we only kept observations from the preliminary control sample with known values for the four indicators (which represent 57,256 observations for the USPTO and 32,952 observations for the EPO control samples).

The matching method used was then “nearest neighbor matching” (without replacement), meaning that each observation from the treated sample was compared to the closest neighbor from the matched control sample, based on their propensity scores.

---

<sup>2</sup> <https://www.accessdata.fda.gov/scripts/cder/ob/index.cfm>.

Once the propensity scores are computed and the matching is completed, we derived the average treatment effect using linear regressions for each of the four outcome variables (size of the family, claims, renewal, and originality).

Descriptive statistics regarding the propensity score matching are available in the *Supplementary material*, including the first stage model predicting the probability of being patents citing FRM-funded publications (B1), the density of propensity scores by subsamples (B2), and the distribution of samples before and after matching against the selected covariates (B3). The results of the propensity score matching are included in *Section 4.2*.

Finally, as a robustness test, we also applied another method, the raking ratio method (Daza, 2012; Deville & Särndal, 1992; Deville, Sarndal, & Sautory, 1993), to ensure that the results of the two samples were comparable. We applied the method of adjusting sample weights to the same three variables used for the propensity score matching: filing year, technology class IPC4, and institutional sector of the applicant. The advantage of the propensity score matching method is that it allows computing a significance test, which the raking ratio method does not. Note that all quality indicators were selected keeping in mind the small size and age of the sample, preventing the use of indicators based on forward patents. The results of the raking method are presented in the *Supplementary material*.

### 3.6. Control Sample of the Publications Funded by FRM

We also used the raking ratio method to construct a control sample of publications in order to examine whether the results obtained for the funding agency (FRM) are generalizable for other institutions with the same characteristics.

Thus, we first constructed a sample identical to that of the FRM corpus based on the following variables:

- a flag that designates whether the publication has been cited in a patent or not;
- a flag that designates whether the publication is open access or not;
- a flag indicating whether the publication has received funding from the ERC;
- a flag indicating whether the publication has at least one European address (EU27);
- five classes that designate the number of countries (1, 2, 3, 4, or 5 and more);
- five classes that designate the number of funding sources for the publication—based on WoS acknowledgment information—(0, 1, 2, 3, or 4 and more);
- dummies by discipline (OST classification in 27 ERC panels); and
- the publication year.

Then, we applied the method to compute the raking weights.

## 4. RESULTS

In this section, we first present an overview of the NPRs in the WoS database. Second, we analyze the average number of citations normalized according to the double NPR/funding crossover for all the publications, as well as for the FRM publications. Finally, we analyze the patents quality indicators of NPR.

### 4.1. NPR, Research Funding and Citation Impact

Table 1 shows that NPRs represent 3% of publications in Natural Sciences and Engineering (NSE) in the WoS database. Of these, 2% received funding (from a project). In addition, over

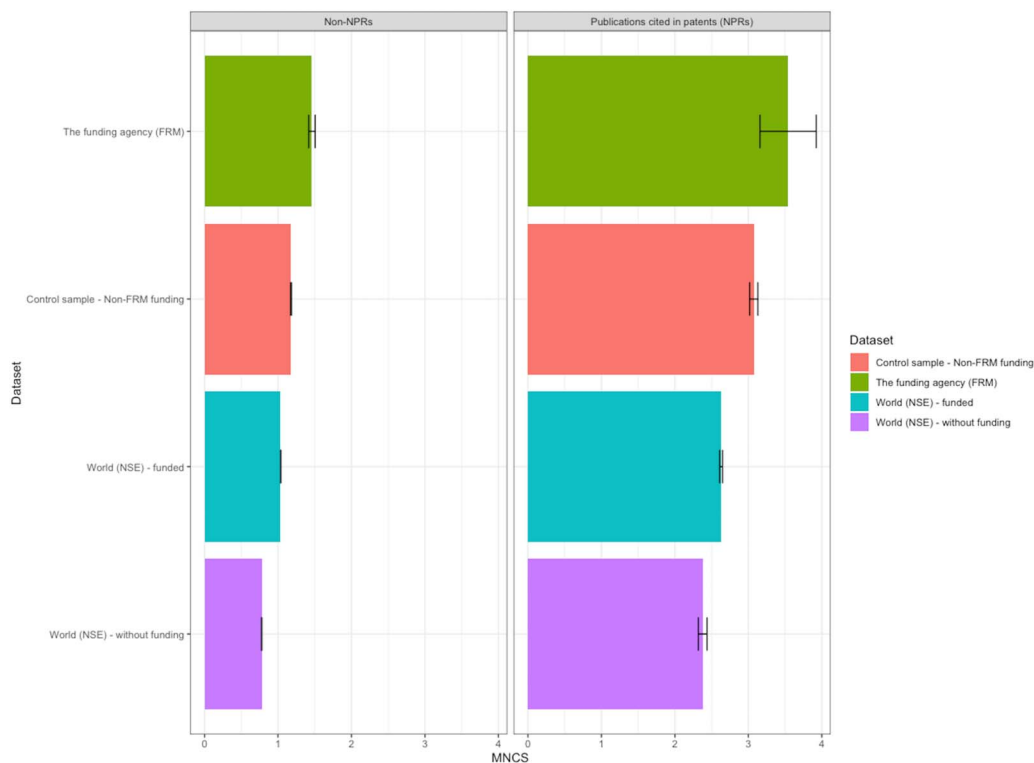


**Table 1.** Distribution of publications according to whether they are NPRs and/or have received funding in Natural Sciences and Engineering (NSE) in the WoS database

Is NPR	Is funded	# Publications in WoS	% Publications in WoS
Yes	Yes	230,165	2%
Yes	No	84,981	1%
No	Yes	6,824,915	60%
No	No	4,201,376	37%
Total (NSE)		11,341,437	100%

the period 2009–2018, nearly two-thirds of NSE publications received dedicated funding. The share of publications not cited in patents and not having received funding is 37%. Citation in a patent is therefore generally rare and strongly linked to the fact that the research has received specific funding (see the Supplementary material for more information about NPRs in the WoS database).

Figure 2 shows the impact of publications’ MNCS (for the calculation method see Leydesdorff and Opthof [2010]) whether they are cited in patents or not, by data set. Figure 2 presents a strikingly clear hierarchy between funded and nonfunded publications. Funded publications receive significantly more citations than their nonfunded counterparts. What is particularly noteworthy in this graph is the distinction between NPRs and non-NPRs. It is evident that NPRs are substantially more cited than non-NPRs, regardless of the data set under consideration. Notably, both FRM-funded publications cited in patents and those in the control group



**Figure 2.** Field normalized impact by set.

(comprising publications similar to FRM-funded ones but funded by other agencies) exhibit citation rates significantly higher than the global average (represented by the WoS database). Specifically, the MNCS for FRM-funded publications is 3.52 compared to 2.63 for the global data set, indicating a 34% higher MNCS for FRM-funded publications. In a broader context, publications cited in patents within the WoS database are 2.55 and 3.12 times more cited than their nonpatent-cited counterparts for funded and nonfunded publications, respectively. These findings underscore the substantial impact of patent citations on the scholarly recognition of research, with NPRs playing a particularly influential role in this dynamic.

Furthermore, it is important to note that the distinction between NPRs from EPO and USPTO has revealed a notable disparity in the impact of publications cited. The findings (not reported here for more simplicity) underscore the distinctive nature of patent citations in the two jurisdictions. Notably, publications cited in EPO patents appear to wield a greater influence, suggesting their closer alignment with the inventive process and practical applications. On the other hand, USPTO patents tend to draw from a broader spectrum of scientific literature, often reflecting the state of the art.

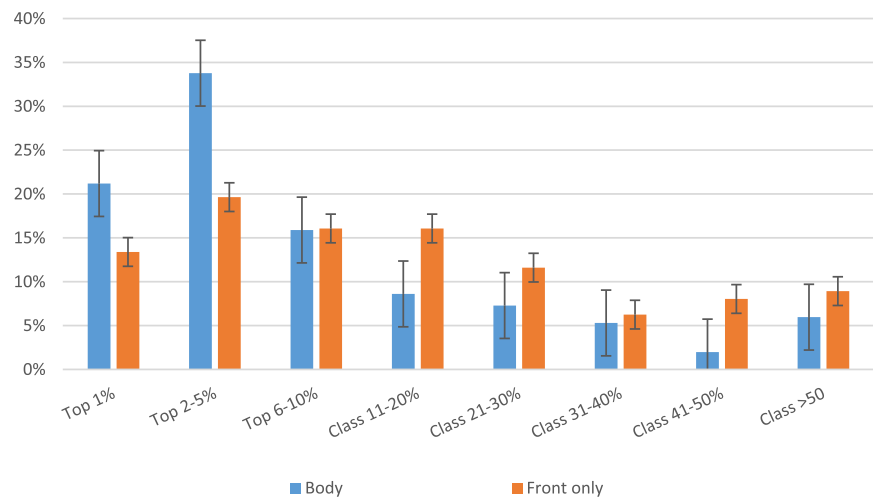
Table 2 shows the distribution of FRM-funded publications according to citation classes (scientific excellence) by type (NPRs and non-NPRs). The table shows that 3% of publications not cited in patents are in the most cited percentile (three times the world average). The rate is five times higher for NPRs with 16% of publications in the most cited percentile. The rate is also high for percentiles between 2% and 10%. The rate is 10% in the 2–5% percentiles for non-NPRs, compared to 26% for NPRs. For percentiles from 6% to 10%, the rate is 10% for non-NPRs and 17% for NPRs.

Overall, the rate of publications in the most cited decile is 23% for non-NPRs, compared to nearly 60% for NPRs.

Figure 3 shows the distribution of FRM publications by scientific excellence class, according to the positioning of the NPRs either on the first page of the patent document or in the body of the text. These are the citations indexed in the U.S. database. The figure highlights some interesting results. We observe that the publication share in the percentile of the most cited publications is significantly higher for the publications (NPRs) cited in the body of the text (21% for the NPRs in the body of the text against 13% for the NPRs in the front page). The

**Table 2.** Distribution of FRM-funded publications by citation class (excellence), 2009–2019

Citation class	# Non-NPRs	# NPRs	% Non-NPRs	% NPRs
Top 1%	183	52	3	16
Top 2–5%	650	83	10	26
Top 6–10%	704	54	10	17
Class 11–20%	1081	38	16	12
Class 21–30%	948	33	14	10
Class 31–40%	703	19	10	6
Class 41–50%	695	15	10	5
Class >50	1,867	22	27	7
<b>Total</b>	<b>6,831</b>	<b>316</b>	<b>100</b>	<b>100</b>



**Figure 3.** Distribution of FRM publications by scientific excellence class according to the positioning of the NPR: first page of the patent document or body of the text (US office).

same goes for the top 2 at 5% most cited class, with a share of 35% for the NPRs in the body of the text, against nearly 20% for the NPRs on the front page.

Citation practices in the USPTO are different from those in other offices (such as the EPO). Thus, patent applicants in the United States are required to review the state of the art in the field and cite on the front page all the works that refer to it, even if they were not used for the invention. Front-page citations are often added by examiners. Bryan et al. (2020) showed that the publications cited in the body of the text are closer to the invention than those on the front page. Moreover, the result presented in Figure 3 suggests that the NPRs in the body of the text have a better impact. We can therefore assume that the closer a scientific publication is to the technology, the more it will have a high citation impact. To confirm this hypothesis would require a more in-depth and systematic analysis.

#### 4.2. Patents Citing NPRs from the FRM Data Set

From 2004 to 2018, 648 patents filed at the USPTO and 110 patent applications filed at the EPO include at least one citation of a publication funded by the FRM, either on the front page or in the body text of the patent.

These patents are mainly from Pharmaceuticals, Biotechnology, and Analysis of Biological Materials technologies (around 85% of patents), according to the WIPO classification in 35 technology fields (Schmoch, 2008) (see Supplementary material). At the USPTO, citing patents are mainly from Pharmaceutical Technologies (70% of the total) and to a lesser extent, from Analysis of Biological Material Technologies (9%), at the EPO the importance of Biotechnology appears to be higher (29%), closely following Pharmaceutical Technologies (43%). The analysis of specialization indexes in the Supplementary material confirms the relative importance of Pharmaceuticals in the USPTO sample compared to the EPO sample (the opposite applies for Biotechnology).

Applicants citing the FRM differ between the USPTO and EPO. At the USPTO, applicants that cite publications funded by FRM are mainly firms (50% of patents), and nonprofit institutions represent 31% and individuals 19% of the patents. At the EPO, on the contrary, applicants citing publications funded by FRM are mainly nonprofit institutions (63%), and firms represent 36% of applications (see Supplementary material). This sectoral difference between

applicants may also explain partially the technology specificities of both samples (nonprofit institutions, universities in particular, are usually more specialized in biotechnologies whereas businesses are more specialized in pharmaceutical technologies).

Finally, inventors of patents citing FRM are also different at the USPTO and EPO. At the USPTO, 61% of them are from the United States (above their average at the office of 47%) and only 11% are from France and 6% from Germany (see Supplementary material). At the EPO, 38% of inventors are from France (6% on average at the office), and only 17% of them are from the United States (26% on average at the office).

These findings suggest that patents citing FRM publications are institutionally and geographically closer to the funded publications data set at the EPO than at the USPTO.

The quality of patents citing the FRM was then investigated. To ensure that our samples at the EPO and USPTO are compared with equivalent control samples, we used two methods: propensity score matching and the raking ratio method. See Section 3.5 for more information about the methodology.

In terms of output variables, we selected four quality indicators: the size of the family, the number of claims, renewals, and the originality index. These indicators were selected from the OECD database of quality indicators and are described in their “Measuring Patent Quality” report (Squicciarini, Dernis, & Criscuolo, 2013).

As explained in more detail in Section 3.5, we used propensity score matching to compare each FRM observation (the treated sample) against the nearest neighbor observation in the control sample based on a set of covariates (application year, technology class, and institutional sector; see the Supplementary material for descriptive statistics on propensity score matching). Then, using a linear regression, we derived the average treatment effect of the FRM sample against the matched control sample. The outcome analysis showed that the FRM data set at the USPTO presents three out of four quality indicators (Table 3) that are significantly higher than those of its matched control sample (number claims, renewals, and originality). In contrast, the average family size of the FRM sample does not appear to be significantly different.

As a robustness test, we also conducted the analysis by using the raking ratio method. Using this method, the four indicators appear to be higher in the FRM sample, with a higher difference for the family size and claims (see Supplementary material).

If we cannot explain why the family size appears to be higher only with the raking ratio method, the results seem to indicate overall that patents citing FRM-funded publications at the USPTO are relatively high-quality patents.

**Table 3.** Average treatment effect of patents citing FRM, USPTO, 2006–2018

Explicative variables	Family size		Claims		Renewal		Originality	
	Coefficient	Pr (>  z )	Coefficient	Pr (>  z )	Coefficient	Pr (>  z )	Coefficient	Pr (>  z )
y-intercept	10.13***	<2.0 × 10 <sup>-16</sup>	14.42***	<2.0 × 10 <sup>-16</sup>	2.40***	<2.0 × 10 <sup>-16</sup>	0.85***	<2.0 × 10 <sup>-16</sup>
FRM	-0.61	0.20	3.80***	4.6 × 10 <sup>-16</sup>	0.65***	<2.0 × 10 <sup>-16</sup>	0.022**	0.022

\*\*\* Significant at 0.1%.

\*\* Significant at 1%.

\* Significant at 5%.

**Table 4.** Average treatment effect of patents citing FRM, EPO, 2004–2018

Explicative variables	Family size		Claims		Renewal		Originality	
	Coefficient	Pr (>  z )	Coefficient	Pr (>  z )	Coefficient	Pr (>  z )	Coefficient	Pr (>  z )
y-intercept	4.42***	$<2.0 \times 10^{-16}$	15.03***	$<2.0 \times 10^{-16}$	3.44***	$<2.0 \times 10^{-16}$	0.70***	$<2.0 \times 10^{-16}$
FRM	0.71	0.26	-1.03	0.27	0.14	0.70	0.083**	0.0095

\*\*\* Significant at 0.1%.

\*\* Significant at 1%.

\* Significant at 5%.

The FRM data set at the EPO is much smaller (105 observations) and presents three out of four indicators with nonsignificant coefficients in comparison with the control sample, using the propensity score matching and linear regression method (family size, claims, and renewals). The only variable that appears to be significantly higher is originality (Table 4).

Similarly, by using the raking ratio method (see Supplementary material), also three out of four variables appear to have relatively similar value in comparison with the control sample. In contrast, the average number of renewals appears to be lower than in the control sample. Please note that not all observations in the EPO sample have a known (i.e., nonempty) value for the output. Although three out of four variables have at least 89 observations with a known value, renewal value is known for only 55 observations, which might explain why the measurement is more challenging for this variable.

Overall, both methods seem to indicate that the patents citing FRM at EPO are relatively similar patents to comparable patents in terms of quality.

This difference between USPTO and EPO patents in terms of quality indicators appears to be surprising at first. One reason for this might come from the institutional sector of EPO patents citing FRM, which is more oriented towards the nonprofit sector than USPTO patents. Traditionally, this sector is usually closer to science. By applying propensity score matching or the raking ratio method, the adjusted control sample also gives a high weight to the nonprofit sector, which is consequently likely to be also close to science. The EPO sample is also of very small size (between 55 and 105 observations depending on the outcome variable analyzed), which also slightly weakens the interpretation of these results. It would be interesting to replicate this analysis in the future with a larger sample.

Finally, five USPTO patents citing the FRM-funded publications have led to drug products approved by the FDA according to the Orange Book database. These drugs come from the private sector, and in particular from two companies: ARRAY BIOPHARMA (now in the Pfizer group) and LOXO ONCOLOGY.

## 5. CONCLUSION AND DISCUSSION

The aim pursued in this study is to contribute to improving the understanding of the contribution of science on technology, using scientific citations in patent documents. This work uses a data set from the FRM as a case study.

Our findings indicate that NPRs account for 3% of the Natural Sciences and Engineering publications in the WoS database, which is used as a proxy for global publications. The rate is higher for the FRM publication sample, with 4% of the publications cited in patents.

The distribution of NPRs by country partly reflects their technological specificities, in particular the number of patents filed by each country, as the countries tend to cite more national scientific publications in their patents. For example, the United Kingdom comes in fifth position for the share of NPRs, behind Germany and Japan, but it is third for all publications (OST, 2019), because the United Kingdom files fewer patents each year than these two countries.

In terms of disciplinary distribution, the areas with the highest share of NPR are Molecular and Structural Biology and Biochemistry, Cellular and Developmental Biology (5% for the world and the FRM data set), Immunity and Infection (3.8% worldwide and 5% FRM data set).

The most striking result is that publications cited in patents are highly cited by other scholarly publications. Our results indicate that publications that are both funded and cited in patents are cited by other publications 2.6 times more than the global publication average. The rate is 3.5 for the NPRs in the FRM data set, an impact 25% greater than all of the NPRs in patents. The impact of publications funded but not cited in patents is significantly lower, but still exceeds the average impact of global publications: It is 1.2 for the world, against 1.8 in the FRM data set. These results suggest on the one hand that the quality of the publications in the FRM data set is better than the world average for the same type of publications and, on the other hand, that the publications cited in patents are generally more cited in scientific publications. The same results are obtained by comparing the profiles of the citation classes: 60% of the NPRs in the FRM data set are in the decile of the most cited publications, against 23% for the non-NPR publications.

However, the analysis conducted does not determine why, on average, the articles cited in patents are high scientific-impact articles. Is it because these articles are of high relevance to technology, because of their high quality or originality, or is it because they are famous because of their high impact and therefore are more cited in the literature? The answer might be a mix of the two elements. Other methods could help to explore this topic further in future work, such as natural language processing methods.

Another interesting result is that scientific publications cited in the body of the patent document are cited more than those cited on the front page. As shown in Bryan et al. (2020), NPRs integrated into the body of the text are closer to the content of the invention. This suggests that publications close to the invention are cited more on average.

The patents filed at the USPTO that cite publications from the FRM data set are mostly filed by the private sector, and their inventors come from various countries outside France (the majority from the United States). These patents also exhibit quality indicators that are superior to the indicators of an adjusted control sample. This last result supports the hypothesis of the existence of more depth of technology, as it is closer to science than to the market. Studies carried out in the field of business innovation find that cooperation with public research is also closer to science and, in general, that these patents have quality indicators that are higher than the general average (Lissoni & Montobbio, 2015). Moreover, the applications filed at the EPO citing publications from the FRM are geographically and institutionally different from citing patents at the USPTO (the inventors are mostly from France and the public sector) and exhibit similar quality indicators, on a smaller data set, in comparison with an adjusted control sample.

By highlighting the positive impact and visibility that research with potential utility for innovation can have, our findings could encourage institutions and researchers to (re)value research that is more likely to have a technological impact. This could ultimately contribute to the development of policies that promote the translation of scientific knowledge into innovations that benefit society (while acknowledging that not all research that benefits society is necessarily applied).



We also advocate for the use of scientific citations in patents in the evaluations of public research organizations and funding agencies. This would potentially significantly improve the understanding of the contribution of publicly funded science on technology, and on the impact on society as a whole.

Finally, in future research, we could consider a deeper analysis taking several issues into account. The projects to which the FRM contributes are generally cofunded, and it would be interesting to analyze the interactions between funding sources, in particular to assess the impact of the project selection processes. It would also be interesting to break down the analysis by project, or to characterize through a semantic analysis the nature and themes of the publications financed by the FRM and their trends over time.

#### ACKNOWLEDGMENTS

The authors would like to thank the two reviewers for their valuable comments, which have significantly contributed to improving the quality of this paper. The authors warmly thank Frédérique Sachwald for her careful proofreading and her remarks, which significantly improved the quality of the paper. The authors would also like to thank Matt Marx and Aaron Fuegi for providing the patent data used in this study.

#### AUTHOR CONTRIBUTIONS

Justin Quemener: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing. Luis Miotti: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing—review and editing. Abdelghani Maddi: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing—original draft preparation, Writing—review and editing.

#### COMPETING INTERESTS

The authors have no competing interests.

#### FUNDING INFORMATION

The authors would like to thank the French Foundation for Medical Research (FRM) for funding the research project that led to this publication.

#### DATA AVAILABILITY

The publication data cannot be made available as it is proprietary to Clarivate Analytics. However, information on patents citing these publications is available here: <https://zenodo.org/records/5111261>.

#### REFERENCES

- Ahmadpoor, M., & Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science*, *357*(6351), 583–587. <https://doi.org/10.1126/science.aam9527>, PubMed: 28798128
- Álvarez-Bornstein, B., & Bordons, M. (2021). Is funding related to higher research impact? Exploring its relationship and the mediating role of collaboration in several disciplines. *Journal of Informetrics*, *15*(1), 101102. <https://doi.org/10.1016/j.joi.2020.101102>
- Bikard, M., & Marx, M. (2020). Bridging academia and industry: How geographic hubs connect university science and corporate technology. *Management Science*, *66*(8), 3425–3443. <https://doi.org/10.1287/mnsc.2019.3385>
- Bryan, K. A., Ozcan, Y., & Sampat, B. (2020). In-text patent citations: A user's guide. *Research Policy*, *49*(4), 103946. <https://doi.org/10.1016/j.respol.2020.103946>

- Carpenter, M. P., & Narin, F. (1983). Validation study: Patent citations as indicators of science and foreign dependence. *World Patent Information*, 5(3), 180–185. [https://doi.org/10.1016/0172-2190\(83\)90139-4](https://doi.org/10.1016/0172-2190(83)90139-4)
- Chen, J.-R., Kan, K., & Tung, I.-H. (2016). Scientific linkages and firm productivity: Panel data evidence from Taiwanese electronics firms. *Research Policy*, 45(7), 1449–1459. <https://doi.org/10.1016/j.respol.2016.03.023>
- Corsini, A., & Pezzoni, M. (2023). Does grant funding foster research impact? Evidence from France. *Journal of Informetrics*, 17(4), 101448. <https://doi.org/10.1016/j.joi.2023.101448>
- Cross, S., Rho, Y., Reddy, H., Pepperrell, T., Rodgers, F., ... Keestra, S. (2021). Who funded the research behind the Oxford–AstraZeneca COVID-19 vaccine? *BMJ Global Health*, 6(12), e007321. <https://doi.org/10.1136/bmjgh-2021-007321>, PubMed: 34937701
- Daza, S. (2012). Raking weights with R. *Sebastian Daza*. <https://sdaza.com/blog/2012/raking/>
- Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84(1), 151–161. <https://doi.org/10.1162/003465302317331982>
- Deville, J.-C., & Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87(418), 376–382. <https://doi.org/10.1080/01621459.1992.10475217>
- Deville, J.-C., Sarndal, C.-E., & Sautory, O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, 88(423), 1013–1020. <https://doi.org/10.1080/01621459.1993.10476369>
- Du, J., Li, P., Guo, Q., & Tang, X. (2019). Measuring the knowledge translation and convergence in pharmaceutical innovation by funding-science-technology-innovation linkages analysis. *Journal of Informetrics*, 13(1), 132–148. <https://doi.org/10.1016/j.joi.2018.12.004>
- Fleming, L., Greene, H., Li, G., Marx, M., & Yao, D. (2019). Government-funded research increasingly fuels innovation. *Science*, 364(6446), 1139–1141. <https://doi.org/10.1126/science.aaw2373>, PubMed: 31221848
- Garg, K. C., Gupta, B. M., Jamal, T., Roy, S., & Kumar, S. (2005). Assessment of impact of AICTE funding on R&D and educational development. *Scientometrics*, 65(2), 151–160. <https://doi.org/10.1007/s11192-005-0264-5>
- Gazni, A., & Ghaseminik, Z. (2019). The increasing dominance of science in the economy: Which nations are successful? *Scientometrics*, 120(3), 1411–1426. <https://doi.org/10.1007/s11192-019-03161-5>
- Hausman, N. (2022). University innovation and local economic growth. *The Review of Economics and Statistics*, 104(4), 718–735. [https://doi.org/10.1162/rest\\_a\\_01027](https://doi.org/10.1162/rest_a_01027)
- Henderson, R., Jaffe, A. B., & Trajtenberg, M. (1998). Universities as a source of commercial technology: A detailed analysis of university patenting, 1965–1988. *The Review of Economics and Statistics*, 80(1), 119–127. <https://doi.org/10.1162/003465398557221>
- Howell, S. T. (2016). *Financing innovation: Evidence from R&D grants* (SSRN Scholarly Paper 2687457). <https://doi.org/10.2139/ssrn.2687457>
- Jowkar, A., Didegah, F., & Gazni, A. (2011). The effect of funding on academic research impact: A case study of Iranian publications. *Aslib Proceedings*, 63(6), 593–602. <https://doi.org/10.1108/00012531111187243>
- Ke, Q. (2020). Technological impact of biomedical research: The role of basicness and novelty. *Research Policy*, 49(7), 104071. <https://doi.org/10.1016/j.respol.2020.104071>
- Leydesdorff, L., Bornmann, L., & Wagner, C. S. (2019). The relative influences of government funding and international collaboration on citation impact. *Journal of the Association for Information Science and Technology*, 70(2), 198–201. <https://doi.org/10.1002/asi.24109>, PubMed: 32728598
- Leydesdorff, L., & Opthof, T. (2010). Remaining problems with the “New Crown Indicator” (MNCS) of the CWTS. *arXiv*. <https://doi.org/10.48550/arXiv.1010.2379>
- Li, D., Azoulay, P., & Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science*, 356(6333), 78–81. <https://doi.org/10.1126/science.aal0010>, PubMed: 28360137
- Lissoni, F., & Montobbio, F. (2015). The ownership of academic patents and their impact. *Revue Economique*, 66(1), 143–171. <https://www.cairn.info/revue-economique-2015-1-page-143.htm>. <https://doi.org/10.3917/reco.661.0143>
- Marx, M., & Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*, 41(9), 1572–1594. <https://doi.org/10.1002/smj.3145>
- Marx, M., & Fuegi, A. (2022). Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations. *Journal of Economics & Management Strategy*, 31(2), 369–392. <https://doi.org/10.1111/jems.12455>
- Mowery, D. C., & Sampat, B. N. (2004). The Bayh-Dole Act of 1980 and university–industry technology transfer: A model for other OECD governments? *The Journal of Technology Transfer*, 30(1), 115–127. <https://doi.org/10.1007/s10961-004-4361-z>
- Narin, F., & Noma, E. (1985). Is technology becoming science? *Scientometrics*, 7(3), 369–381. <https://doi.org/10.1007/BF02017155>
- OST. (2019). *Dynamics of scientific production in the world, in Europe and in France, 2000–2016* (p. 100). HCERES. <https://www.hceres.fr/fr/publications/dynamics-scientific-production-world-europe-and-france-2000-2016-ost>
- Poege, F., Harhoff, D., Gaessler, F., & Baruffaldi, S. (2019). Science quality and the value of inventions. *Science Advances*, 5(12), eaay7323. <https://doi.org/10.1126/sciadv.aay7323>, PubMed: 31844673
- Popp, D. (2017). From science to technology: The value of knowledge from different energy research institutions. *Research Policy*, 46(9), 1580–1594. <https://doi.org/10.1016/j.respol.2017.07.011>
- Roach, M., & Cohen, W. M. (2013). Lens or prism? Patent citations as a measure of knowledge flows from public research. *Management Science*, 59(2), 504–525. <https://doi.org/10.1287/mnsc.1120.1644>, PubMed: 24470690
- Roshani, S., Bagherlooieh, M.-R., Mosleh, M., & Coccia, M. (2021). What is the relationship between research funding and citation-based performance? A comparative analysis between critical disciplines. *Scientometrics*, 126(9), 7859–7874. <https://doi.org/10.1007/s11192-021-04077-9>
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3), 169–188. <https://doi.org/10.1023/A:1020363010465>
- Schmoch, U. (2008). *Concept of a technology classification for country comparisons* (p. 15). World Intellectual Property Organization (WIPO).
- Silva, F., & Carreira, C. (2012). Do financial constraints threaten the innovation process? Evidence from Portuguese firms. *Economics of Innovation and New Technology*, 21(8), 701–736. <https://doi.org/10.1080/10438599.2011.639979>
- Sohn, S. Y., Gyu Joo, Y., & Kyu Han, H. (2007). Structural equation model for the evaluation of national funding on R&D project of SMEs in consideration with MBNQA criteria. *Evaluation and*

- Program Planning*, 30(1), 10–20. <https://doi.org/10.1016/j.evalprogplan.2006.10.002>, PubMed: 17689310
- Squicciarini, M., Dernis, H., & Criscuolo, C. (2013). *Evaluer la qualité des brevets: Des indicateurs de valeur technologique et économique*. <https://doi.org/10.1787/5k4522wkw1r8-en>
- Velayos-Ortega, G., & López-Carreño, R. (2023). Indicators for measuring the impact of scientific citations in patents. *World Patent Information*, 72, 102171. <https://doi.org/10.1016/j.wpi.2023.102171>
- Veugelers, R., & Wang, J. (2019). Scientific novelty and technological impact. *Research Policy*, 48(6), 1362–1372. <https://doi.org/10.1016/j.respol.2019.01.019>
- Wang, J., & Verberne, S. (2021). Two tales of science technology linkage: Patent in-text versus front-page references. *arXiv*. <https://doi.org/10.48550/arXiv.2103.08931>
- Wang, L., & Li, Z. (2021). Knowledge flows from public science to industrial technologies. *The Journal of Technology Transfer*, 46(4), 1232–1255. <https://doi.org/10.1007/s10961-019-09738-9>
- Yan, E., Wu, C., & Song, M. (2018). The funding factor: A cross-disciplinary examination of the association between research funding and citation impact. *Scientometrics*, 115(1), 369–384. <https://doi.org/10.1007/s11192-017-2583-8>