The prevalence of NIH-funded research in commercial products

Authors: Gaétan de Rassenfosse[1*], Emilio Raiteri[2]

Affiliations:

[1] College of Management of Technology, École polytechnique fédérale de Lausanne.

[2] Technische Universiteit Eindhoven.

[*] Correspondence to: Station 5, 1015 Lausanne, Switzerland. gaetan.derassenfosse@epfl.ch

DRAFT. PLEASE DO NOT DISTRIBUTE. COMMENTS WELCOME.

THIS VERSION: JULY 12, 2019.

Abstract:

This paper explores the prevalence of NIH-funded research in commercial products using a novel method. We find that more than half of commercial products by biotechnology and pharmaceutical firms, and 12 percent by medical device firms, exploit NIH-funded research. Products that are commercialized by smaller firms are more likely to involve direct NIH-funded research. Finally, science takes more than 20 years to translate into commercial products. Our method can be adopted to study the role of government funding in products across a range of industries.

Main Text:

Public funding of R&D accounts for about a fifth of overall R&D expenditures in the United States and represents more than $120 billion per year according to the OECD.[1] Government funding plants the seeds for future innovations by helping establish the knowledge base on which private companies build to develop innovative products. But do these seeds actually grow? And how long does blossoming take? Answering these questions is central to the assessment of the 'real impact' of publicly funded research.

U.S. federal agencies, facing the need to maximize the scientific returns on the taxpayers' investments (Lorsch, 2015) and to justify their budget to Congress, have been exploring these questions for a long time. Given the difficulty of tracking basic into the end product, they have resorted primarily to case study methods (e.g., National Academies 2016). Scholars have followed suit, also relying on case studies (e.g., Mazzucato 2015). Regarding funding by the National Institute of Health (NIH) more specifically, scholars have also performed more systematic studie on the role of public funding on FDA-approved drugs using data from the Orange Book (Sampat 2009, Sampat and Lichtenberg 2011, Stevens et al. 2011, Li et al. 2017, Azoulay et al. 2018, Cleary et al. 2018) or using data on NIH licensing agreements (Chatterjee and Rohrbaugh 2014). Finding ways of documenting the role of government funding in a more

[1] Source: "Gross domestic expenditure on R&D by sector of performance and source of funds" available at https://stats.oecd.org/Index.aspx?DataSetCode=GBARD_NABS2007, last accessed July 12, 2019.

systematic manner is one of the key challenges of the emerging field of Science of Science Policy (Lane et al. 2011*, inter alia*).

This study takes a step in this direction. We explore the product portfolios of 140 companies active in the health industry, in particular in the following three sectors: Biotechnology (35 firms), Medical Devices (94), and Pharmaceuticals (11). We link each of the 2968 products in our sample to the 4282 patents protecting them and to the 10,635 scientific publications that these patents cite. We then exploit so-called government interest statements in patent documents and funding acknowledgements in scientific publications to identify research that was sponsored by the National Institute of Health (NIH). Government interest statements signal that the U.S. Government retains rights in inventions that result from federally funded research and development.

The analysis combines data from four main sources. The correspondence between products and patents is obtained from the *IPR*oduct database. It contains data reported on so-called 'virtual patent marking' webpages of company websites (de Rassenfosse 2018). Section 287(a) of 35 U.S.C. (so-called 'marking statute') encourages patentees to provide constructive notice to the public that an article is patented. One way for doing this is to publish a webpage that associates the patented product with the patent number(s). A 'product' in the context of this paper must be understood as a good that is commercially available at the time of data collection. Firms in the *IPR*oduct database were selected at random by crawling the web in search of virtual patent marking webpages. We extract from this database all the firms related the three sectors relevant for the study. Note that the data collected do not overlap with those in the Orange Book—indeed, only 7.4 percent of the patents considered in the present study are listed in the Orange Book (and only 7.5 percent of the patents in the Orange Book are considered in the present stud).

Government interest statements and the link to the contracting federal agencies is obtained from the 3PFL database (de Rassenfosse et al. 2019). The construction of the 3PFL database takes advantage of the U.S. Federal AcquisitionRegulation (FAR). The FAR regulates the federal procurement process and stipulates that federal contractors may retain title to inventions made in the performance of work under a Government contract. When the contractor decides to take title to an invention, it should file a patent application and grant the Government an irrevocable license to use the invention. In that case, the FAR requires the contractor to include in the U.S. patent document a statement acknowledging Government support , together with information about the funding agency and the contract identification number. Regarding research grants, the Bayh-Dole Act imposes requirements similar to FAR for recipients of federally funded research grants. The grantee seeking patent protection for such inventions shall mention the grant number and the agency that issued the grant in the government interest statement.

Finally, data on scientific articles cited in patent documents is scraped from the lens.org website, and the funding sources for these papers is recovered from the *Web of Science* API.

We identify three pathways from NIH funding to commercial products, as illustrated in Fig. 1. The first pathway (P1) relates to products that are protected by (at least) one patent funded by the NIH. The second and third pathways exploit citations made in patent documents. The second pathway (P2) relates to products that are protected by at least one patent that cites at least one patent funded by the NIH. The third pathway (P3) relates to products that are protected by at least one patent that cites a scientific paper funded by the NIH. The documents cited (P2 and P3) need not be published by the commercializing firm. Thee three pathways considered are similar to Fleming et al. (2019).

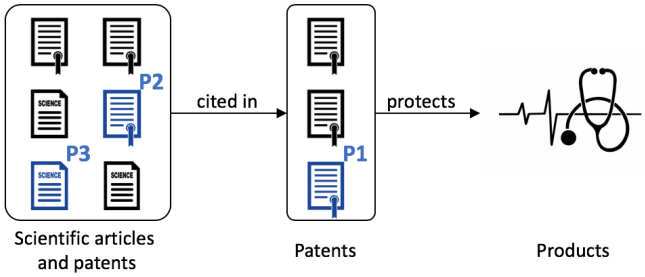

Fig. 1. Pathways from research and development to commercial products

Fig. 2 illustrates the prevalence of each pathway by sector. About 13 percent of products by biotechnology firms in our sample have received direct funding by the NIH (Pathway 1, Panel A) whereas 16 percent of products build indirectly on NIH-funded patents (Pathway 2, Panel B). The impact of NIH funding is particularly important when considering NIH-funded scientific publications (Pathway 3, Panel C). For instance, 45 percent of products by biotechnology firms are protected by a patent that cites at least one NIH-funded scientific article. However, not all patents build on science, and the light-color bars in Panel C considers the subset of products whose patent cites a scientific publication. About 73 percent of biotech products with science-based patents cite a NIH-funded publication.

Notice that the sources of funding can overlap. A product can be protected by a NIH-funded patents, and also cite NIH-funded patents and scientific articles. Panel D of Fig. 2 presents the proportion of products associated with at least one identified pathway. Overall, 53 percent of products by biotechnology firms (directly or indirectly) exploit knowledge that was created thanks to funding by the NIH. The figure reaches 61 percent for products by pharmaceutical firms and 12 percent for products by medical device firms. These estimates are necessarily lower bound estimates of the true rates; they are affect by lack of reporting and misreporting of government interest statements and funding acknowledgements.

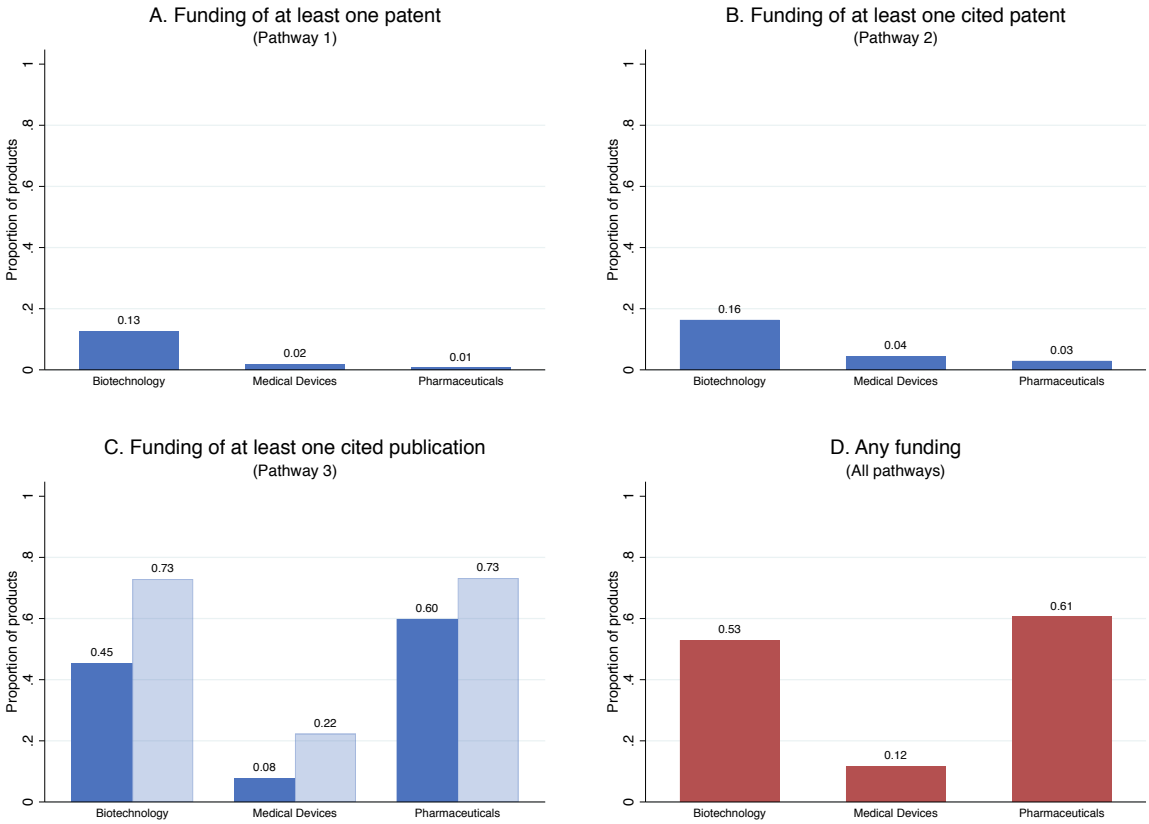

Fig 2. Prevalence of NIH-funded research in commercial products, by pathway

The propensity to exploit NIH-funded research may depend on the characteristics of the firm. Fig. 3 explores differences with respect to the size of the commercializing firm. It reports the marginal effects—associated with different levels of number of employees—of the probability that a product builds on NIH-funded research. The marginal effects are recovered from a probit regression model that controls for the size of the patent portfolio and the activity sector, among other variables.

The results presented in Panel A of Fig. 3 suggest that products by small firms are more likely to embed research findings obtained thanks to direct NIH funding (Pathway 1) than products by large firms. The probability score increases by 31.6 percentage points for firms with 1–10 employees and by 21.4 percentage points for firms with 11–50 employees. Panel B depicts the marginal effects of the probability that a product is protected by a patent that cites NIH-funded research (Pathway 3). The probability score steadily decreases from very small firms (50.7 percentage points for firms with 2–10 employees) to medium-sized firms (22.4 percentage points for firms with 201–500 employees). It then oscillates around 50 percent for larger firms.

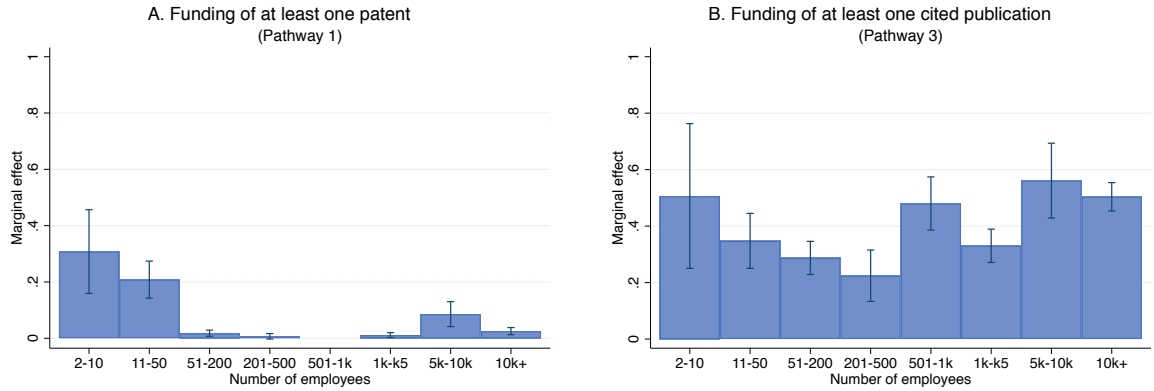

Fig 3. Prevalence of NIH-funded research by pathway and firm size

We next investigate potential differences in the 'translation time'—sometimes referred to as 'gestation lag' (Li and Hall, 2019)—of NIH-funded research. Fig. 4 depicts the frequency distribution of patent age (Pathway 1, Panel A) and cited article age (Pathway 3, Panel B) for NIH-funded research vs. non-NIH-funded research. The mean age of NIH-funded patents protecting commercial products is 10.35 years and the difference with non-NIH patents is not statistically significant (p-value = 0.39). Note that the legal validity of patents is 20 years (with possible exceptions for FDA-approved patented products), which is therefore the upper limit of the distribution. Of course, older technologies may well be embedded into these products.

Overall, the mean age of scientific articles is 23 years (not reported). That is, today's products embed science that was produced more than two decades ago. The oldest cited paper in the sample is 151 years old. It was published on July 27, 1867, in *The Lancet* by Lauder Brunton and is entitled "On the Use of Nitrite of Amyl in Angina Pectoris." It is cited in US patent #7803838B2 granted on September 28, 2010, and protecting the BYVALSON® drug by company Allergan.

The funding acknowledgement dataset does not go as far back in time. Its coverage starts after 1980. Panel B focuses on scientific articles published since 1980. The mean age of NIH-funded articles in this time period is 20.7 years and the difference with non-NIH-funded scientific articles is statistically significant (p-value = 0.0022) but small in magnitude. NIH-funded articles are on average 7 months younger. Thus, overall, we find little differences in the translation time of research between publicly-funded and privately-funded research.[2]

To answer the questions raised at the beginning of the article about whether the seeds grow, the answer is yes, the seeds do grow. They seem to grow more in smaller firms, but firms of all sizes rely on NIH-funded research to a significant extent. More than half of commercial products by biotechnology and pharmaceutical firms exploit NIH-funded research in one way or another. However, blossoming takes time. It takes more than 20 years on average for scientific papers to find their way to the market.

This article demonstrates that it is possible to reconstruct the innovation chain, from original research funding to the final product, on a large scale using available data. However, it is silent on the effectiveness of funding. Answering this question would require to observe the proportion of NIH-funded research that finds its way to the market. This is a pressing question that should be answered. The major hurdle to get there is lack of data. In this regard, the resolution proposed in the recent Civil Society Open Letter to World Health Assembly

[2] Strictly speaking, non-NIH funded research might by funded by other public sources. However, NIH is by far the largest source of public funding for research in the health industry.

Delegates on Transparency Resolution Negotiations would be an important step in the right direction. It would give the World Health Organization and national governments strong mandates to collect and analyze data on drug prices, R&D costs, clinical trial results and costs, patent landscapes, and more.[3]

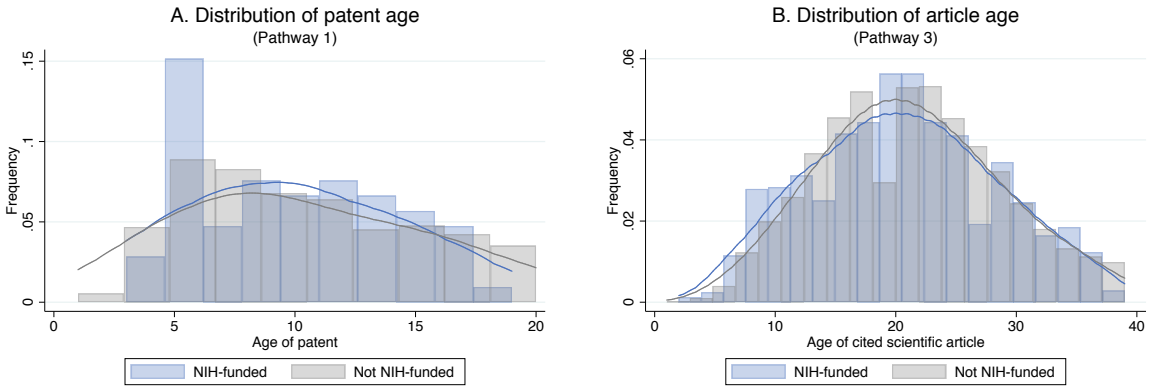

Fig 4. Gestation lag of science and technology

Acknowledgements

The authors are grateful to David Portabella for his work on the *IPR*oduct database. Cyril Verluise also contributed to data collection. The construction of the *IPR*oduct database was supported by grants from the U.S. NSF (EAGER 1645264), the European Patent Office and the Swiss Federal Institute of Intellectual Property. The construction of the 3PFL database received the support of the Marie Skłodowska-Curie COFUND programme at EPFL. The data analysis presented in this paper was supported by a grant from the NBER "Advancing the Science of Science Funding" research program.

References

Azoulay, P., Graff Zivin, J. S., Li, D., & Sampat, B. N. (2018). Public R&D investments and private-sector patenting: evidence from NIH funding rules. *The Review of Economic Studies*, 86(1), 117–152.

Chatterjee, S. K., & Rohrbaugh, M. L. (2014). NIH inventions translate into drugs and biologics with high public health impact. *Nature Biotechnology*, 32(1), 52.

Cleary, E. G., Beierlein, J. M., Khanuja, N. S., McNamee, L. M., & Ledley, F. D. (2018). Contribution of NIH funding to new drug approvals 2010–2016. *Proceedings of the National Academy of Sciences*, 115(10), 2329–2334.

de Rassenfosse, G. (2018). Notice failure revisited: Evidence on the use of virtual patent marking. *National Bureau of Economic Research Working Paper* 24288.

de Rassenfosse, G., Jaffe, A., Raiteri, E. (2019). The procurement of innovation by the U.S. Government. *PLOS ONE*, in press.

[3] Source : "Civil Society Open Letter to World Health Assembly Delegates on Transparency Resolution Negotiations" available at https://www.keionline.org/30701, last accessed July 12, 2019.

Fleming, L., Greene, H., Li, G., Marx, M., & Yao, D. (2019). Government-funded research increasingly fuels innovation. *Science*, 364(6446), 1139–1141.

Lane, J. I., Fealing, K. H., Marburger III, J. H., & Shipp, S. S. (Eds.). (2011). *The science of science policy: A handbook*. Stanford University Press, Stanford, California.

Li, D., Azoulay, P., & Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science*, 356(6333), 78–81.

Li, W., & Hall, B. (2019). Depreciation of business R&D capital. *Review of Income and Wealth*, in press.

Lorsch, J. R. (2015). Maximizing the return on taxpayers' investments in fundamental biomedical research. *Molecular Biology of the Cell*, 26(9), 1578–1582.

Mazzucato, M. (2015). *The entrepreneurial state: Debunking public vs. private sector myths*. Anthem Press, London, Great Britain.

National Academies of Sciences, Engineering, and Medicine. (2016). *SBIR at the National Science Foundation*. National Academies Press.

Sampat, B. N. (2009). Academic patents and access to medicines in developing countries. *American Journal of Public Health*, 99(1), 9–17.

Sampat, B. N., & Lichtenberg, F. R. (2011). What are the respective roles of the public and private sectors in pharmaceutical innovation? *Health Affairs*, 30(2), 332–339.

Stevens, A. J., Jensen, J. J., Wyller, K., Kilgore, P. C., Chatterjee, S., & Rohrbaugh, M. L. (2011). The role of public-sector research in the discovery of drugs and vaccines. *New England Journal of Medicine*, 364(6), 535–541.