Universiteit Leiden
The Netherlands

Novel science for novel technology?
Reinhilde, Veugelers; Jian, Wang

Citation
Reinhilde, V., & Jian, W. (2018). Novel science for novel technology? *Sti 2018 Conference Proceedings*, 807-813. Retrieved from https://hdl.handle.net/1887/66276

Version: Not Applicable (or Unknown)
License: Leiden University Non-exclusive license
Downloaded from: https://hdl.handle.net/1887/66276

Note: To cite this publication please use the final published version (if applicable).

STI 2018 Leiden

23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

© of the text: the authors
© 2018 Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands

This ARTICLE is licensed under a Creative Commons Atribution-NonCommercial-NonDetivates 4.0 International Licensed

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

Novel science for novel technology?[1]

Reinhilde Veugelers* and Jian Wang**

**reinhilde.veugelers@kuleuven.be*
KU Leuven (Belgium), Bruegel (Belgium), and CEPR (UK)

***j.wang@sbb.leidenuniv.nl*
Leiden University (Netherlands) and KU Leuven (Belgium)

Introduction

How well science and industry are interconnected and how scientific knowledge can feed into technology development is nowadays recognized as crucial for the innovative performance, growth and competitiveness of nations (Freeman, 1987; Jaffe, 1989; Nelson, 1993). Corporations are increasingly employing open innovation strategies (Laursen & Salter, 2006) and have been increasingly interested in leveraging public science as an external knowledge source for their technological development (Cockburn & Henderson, 1996; Gambardella; Mansfield, 1998). From the science side, universities and public research organizations have been called upon to engage more actively and directly in knowledge transfer activities and to add entrepreneurial objectives to their missions (Branscomb, Kodama, & Florida, 1999; Etzkowitz & Leydesdorff, 2000; Geuna & Muscio, 2009).

Although many important examples can be given where science was important for initiating new technological applications, there is also widespread perception that most science rarely pays off in generating practical applications (Gittelman & Kogut, 2003). Research still has to uncover which kinds of science and which mechanisms/processes generate more and more effective links between science and industry. Most of the literature on industry-science-links takes the firms as the starting point and examines which firms benefit from linking to science and how (Cassiman, Veugelers, & Zuniga, 2008; Fleming & Sorenson, 2004). Much less developed is the literature taking the science as the starting point and examining which science is most likely to be used in technological development. This latter approach can reveal characteristics of scientific outputs which are in particular relevant for technological innovation and thus add to our understanding of the interplay between science and technology.

In this contribution, we take the science perspective of links between science and technology and aim to identify the scientific contribution which is most likely to be referenced as prior art by patented inventions. When examining characteristics of the science, we are particularly interested in novel research, because of its high gain/high risk profile. Breakthrough science often requires novel approaches, which often faces a higher level of uncertainty. Wang, Veugelers, and Stephan (2017) find that novel scientific publications, identified as those

[1] Financial support from KU Leuven (GOA/12/003) and FWO (G.0825.12 and a postdoctoral fellowship) is gratefully acknowledged. Publication data are sourced from Clarivate Analytics Web of Science Core Collection and patent data from EPO Worldwide Patent Statistical Database - October 2013.

making new combinations of prior knowledge, are more likely to eventually become top cited papers but are also more risky, as reflected in a larger dispersion in their citation distribution. Novel publications also encounter difficulties in getting published in journals with higher Impact Factor and display a delayed recognition in their citation accumulation process. Furthermore, novel publications are highly cited in foreign field but not in their own field.

In addition to the special characteristics of scientific impact for novel research, it is important to study the characteristics of technological impact for novel research. This paper draws on all the Clarivate Analytics Web of Science (WoS) Science Citation Index Expanded (SCIE) journal articles published in 2001 and all the patents in PATSTAT version October 2013. We find that novel scientific articles are significantly more likely to have a technological impact than non-novel publications. Conditional on being cited by patents, novel publication do not display a delayed patent citation process. Furthermore, the technological impact of novel science is significantly unprecedented, reaching more technology fields and new fields previously non-impacted. Finally, novel science is also significantly more likely to lead to technological inventions which are themselves novel.

Data and method

The publication dataset consists of all research articles in SCIE published in 2001, and there are in total 982093 observations. The patent dataset is the ECOOM cleaned EPO, USPTO and WIPO patents of the PATSTAT 201310 version. Non-patent-references in patents are matched to individual publications in WoS, using the algorithm developed by ECOOM (Julie Callaert, Grouwels, & Van Looy, 2012; J. Callaert et al., 2014; Magerman, Van Looy, & Song, 2010).

The focal explanatory variable is scientific novelty. Following Wang et al. (2017), we measure the novelty of a scientific publication as the number of new journal pairs in its references, weighted by the cosine similarity between the newly-paired journals:

$$Novelty = \sum_{J_i - J_j \text{ pair is new}} \left(1 - COS_{i,j}\right)$$

As the distribution of the novelty measure is highly skewed, we use a categorical novelty variable NOV CAT: (1) non-novel, if a paper has no new journal combinations, (2) moderately novel, if a paper makes at least one new combination but has a novelty score lower than the top 1% of its WoS subject category, and (3) highly novel, if a paper has a novelty score among the top 1% of its WoS subject category.

Prior literature observes that highly cited publications are also more likely to be referenced in patents (Ahmadpoor & Jones, 2017; Hicks, Breitzman, Hamilton, & Narin, 2000), so we control for the number of citations received from scientific publications. We also control for the Impact Factor of the journal in which the focal publication appears. Scientific field (i.e., WoS subject category) fixed effects are incorporated to account for field differences. Because of the coverage bias and differences in citation behavior across patent offices, we add three geographic dummies: whether a papers has a (1) US, (2) EPO member state, or (3) Japanese affiliation. To further account for the differences between patent offices, we check the robustness of our results using only USPTO or EPO patent data. In addition, we control for the number of references of the scientific publication and the number of authors, which are found to be correlated with both novelty and impact.

Table 1. Scientific novelty and patent citations.

	(A) Directly cited by patent	(B) Indirect cited by patent	(C) Years of time lag	(D) # citing patents	(E) # citing IPC6s	(F) Cited in new IPC6	(G) Cited by novel patent (sci)	(H) Cited by novel patent (tech)	(I) Cited by novel patent (comb)
	Probit	Probit	OLS	Poisson	Poisson	Probit	Probit	Probit	Probit
NOV CAT2	0.105^{***}	0.062^{***}	-0.082^{**}	0.009	0.012^{+}	0.098^{***}	0.095^{***}	0.041^{**}	0.070^{***}
	(0.006)	(0.005)	(0.026)	(0.017)	(0.007)	(0.017)	(0.016)	(0.014)	(0.019)
NOV CAT3	0.192^{***}	0.116^{***}	-0.105	0.079^{+}	0.031^{+}	0.228^{***}	0.148^{***}	0.086^{*}	0.096^{*}
	(0.017)	(0.015)	(0.071)	(0.044)	(0.017)	(0.043)	(0.040)	(0.038)	(0.049)
Patent citations (ln)					0.594^{***}	0.383^{***}	0.481^{***}	0.664^{***}	0.461^{***}
					(0.003)	(0.007)	(0.006)	(0.006)	(0.007)
N	921193	922141	103378	103378	103289	103282	98752	98757	98668
(Pseudo) R^2	0.174	0.320	0.036	0.093	0.364	0.157	0.133	0.180	0.157

*** $p<.001$, ** $p<.01$, * $p<.05$, $^{+}$ $p<.10$. All regressions incorporate the complete set of control variables: Journal Impact Factor (ln), number of scientific citations (ln), number of co-authors (ln), whether internationally co-authored, number of references (ln), whether have US affiliations, whether have EPO member state affiliations, whether have Japanese affiliations, and WoS subject category dummies. Robust standard errors in parentheses. Data sourced from Clarivate Analytics Web of Science Core Collection and EPO Worldwide Patent Statistical Database - October 2013.

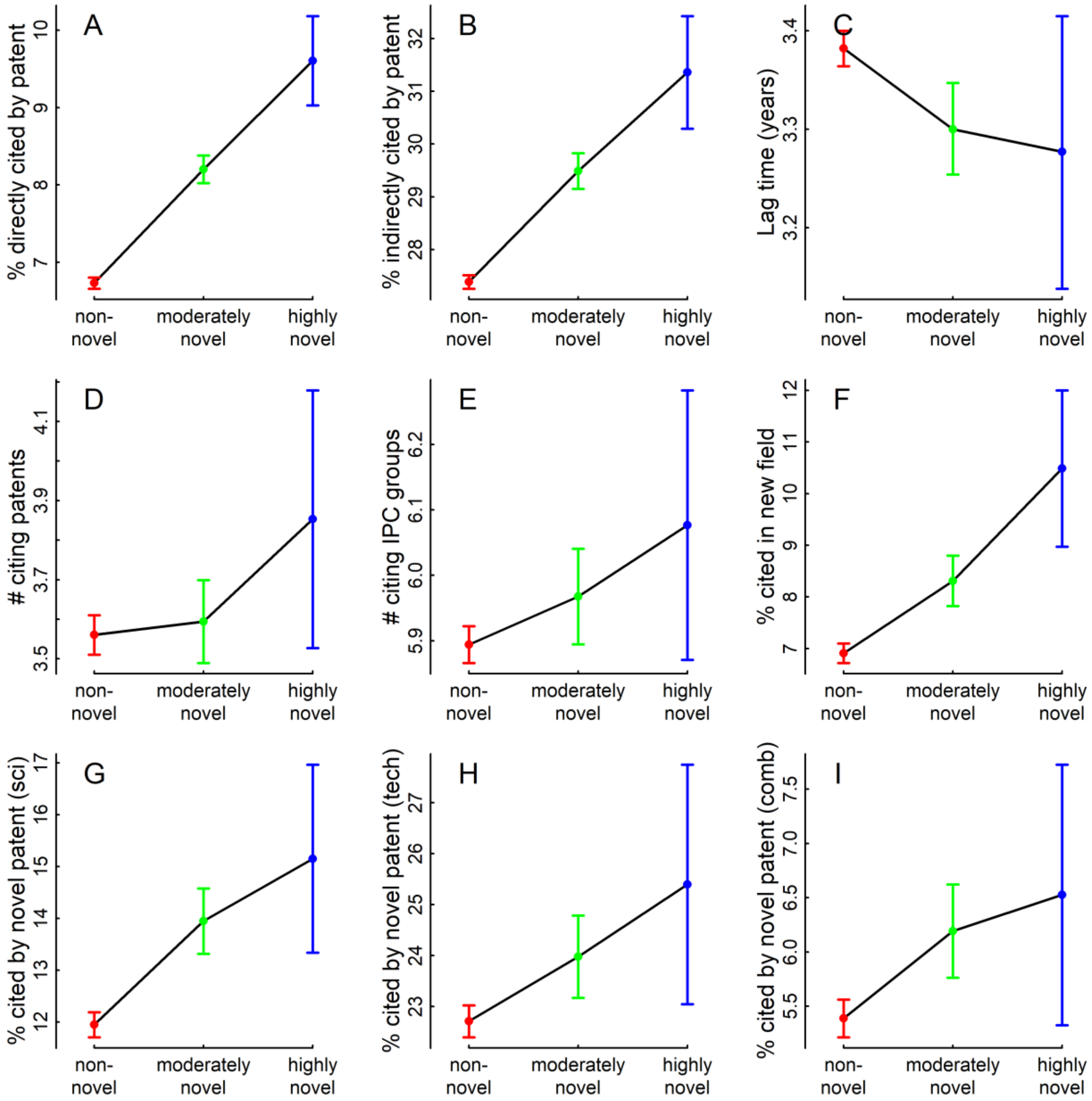
Results

The descriptive statistics give some first indications that novel papers are more likely to have a technological impact than non-novel papers, especially the set of highly novel papers: on average about 10% of the scientific publications in our sample are cited by patents up until 2013, this probability is 16% for highly novel publications (NOV CAT3) and 13% for moderately novel publications (NOV CAT2).

Table 1 reports all the regression results, which are further visualized in Figure 1, where estimates are for an average publication in different novelty categories. Table 1 Column (A) estimates the likelihood of being cited by patents using a Probit model, and the result confirms that novel publications have a higher chance of being cited by patents. The premium for novel science on technological impact is sizeable: calculated at the mean level of all other variables, the probability of being cited by patents is 43% higher for highly novel publications and 22% higher for moderately novel publications, compared with non-novel publications with all else equal. Column (B) additionally investigates the indirect impact, that is, whether a publication is cited by patents indirectly, through another publication. The probability of being indirectly cited by patents is 15% higher for highly novel publications and 8% higher for moderately novel publications, compared with comparable non-novel ones.

In the next step, we scrutinize the characteristics of technological impact, restricting the analysis to the set of publications that are cited by patents. One characteristic of the technological impact that we study is the time lag, where the dependent variable is the difference between the publication year of the scientific article and the application year of the first patent citing the focal scientific article. The result reported in Column (C) shows that the technological impact of novel papers does not face any delay compared with non-novel papers. On the contrary, moderately novel papers take significantly less time to receive their first patent citation than non-novel papers. This contrasts with the results found on the time to scientific impact (Wang et al., 2017), where novel papers display a delayed recognition in scientific citations.

Figure 1: Scientific novelty and patent citations.



Column (D) analyzes the intensity of technological impact, that is, the number of patent citations or the number of patents citing the focal publication. Highly novel publications receive approximately 8% more patent citations than non-novel ones, while there is no significant difference between moderately novel and non-novel publications.

We further examine whether the impact of novel science reaches a broader set of technological domains. We measure the scope of technological impact by the number of IPC groups (6-digit level) in which the scientific publications are cited. Column (E) shows that novel scientific publications are cited in a larger number of technological fields, compared with non-novel publications with the same value in all the control variables and additionally with the same number of patent citations. The technological impact scope premium for novel science is however only weakly significant (10% level) and small in size: 3% higher for highly novel and 1% higher for moderately novel publications.

Column (F) examines whether the technological impact of scientific publications reaches new technological areas, i.e., technological areas that have never cited the scientific field of the focal publication before. The result suggests that novel scientific publications are more likely to have impact in new technological areas than non-novel publications with the same number of patent citations. This premium for novel science is substantial: highly novel publications have a 52% higher probability to be cited in new technology areas, and moderately novel publications have a 20% higher chance, compared with comparable non-novel publications.

Finally, we test whether novel science leads to novel inventions. Following Verhoeven, Bakker, and Veugelers (2016), we identify three types of novel patents: (1) novelty in scientific knowledge origins, (2) novelty in technological knowledge origins, and (3) novelty in recombination. Results in column (G)-(I) show that novel publications are more likely to be referenced as prior art by patents that have themselves novelty features. First, in line with the finding in column (F), novel publications are more likely to be cited by patents with new scientific knowledge origins: highly novel publications are 27% more likely, and moderately novel publications 17%, to be cited by such patents than comparable non-novel publications (Column G). The advantage for novel science to lead to novel patents in terms of new technological knowledge origins is also significant but smaller: highly novel publications are 12%, and moderately novel publications 6%, more likely to be cited by novel patents with new technological knowledge origin, compared with comparable non-novel publications (Column H). Furthermore, novel science is significantly more likely to be referenced by patents that make themselves new combinations of knowledge components as proxied by the combination of technology classes they cover: highly novel patents are 21%, and moderately novel publications 15%, more likely to be cited by novel patents with new combinations (Column I). Thus, novel research is more important as sources for novel inventions.

For robustness checks, we replicate regressions in Column (E) and (F) using IPC subclasses (4-digit level) instead of IPC groups (6-digit level) and obtained robust results. We also run all the regressions (1) using USPTO data only and (2) using EPO data only, results are consistent. In addition, we separated physical sciences and life sciences and observed robust results for both fields.

Conclusion

Drawing on 2001 SCIE journal articles and patents in PATSTAT v201310, we study the relationship between scientific novelty and technological impact, tracing technological impact through patent citations. We find that a handful of novel scientific publications are significantly more likely to have direct and indirect technological impact, in particular the top 1% highly novel scientific publications. Conditional on being directly cited by patents, novel publications do not display a delayed citation process in technological impact, different from their scientific impact. There is a significant but small association between scientific novelty and the intensity of technological impact. More importantly, novel science has a broader technological impact, covering more diverse technological fields and reaching technology fields previously not impacted by the scientific fields of the focal publications. We also found that novel science is more likely to lead to novel inventions.

It is widely accepted that novelty is important for science because of its irreplaceable role in advancing the scientific frontier, and this paper provides further empirical evidence that novel science also has greater technological impact. As there is an increasing pressure on science to be economically and socially relevant, our findings suggest that scientific novelty should be encouraged not only for the sake of scientific progress but also for its greater scope for technological impact, particularly its impact on new areas of technology impact and impact on more novel technology inventions. Therefore, any bias in the current science system against novelty will not only imperil scientific progress but also hinder technological development.

However, any discussion of policy implications should await further robustness checks on the results. Further characteristics of the novel science need to be scrutinized, for example, its interdisciplinary nature, or whether it is starting a new emerging field. In addition, the technology that uses novel science needs to be further examined. Future research should also investigate how the links between novel science and technology inventions play out differently compared with non-novel science: are institutional links, inventor-author personal links, and geographic proximity more important for bring novel science to patented inventions? These are but a few of the many interesting further research questions brought about by this contribution.

References

Ahmadpoor, M., & Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science, 357*(6351), 583-587. doi:10.1126/science.aam9527

Branscomb, L. M., Kodama, F., & Florida, R. L. (1999). *Industrializing knowledge : university-industry linkages in Japan and the United States*. Cambridge, Ma.: MIT Press.

Callaert, J., Grouwels, J., & Van Looy, B. (2012). Delineating the scientific footprint in technology: Identifying scientific publications within non-patent references. *Scientometrics, 91*(2), 383-398. doi:doi:10.1007/s11192-011-0573-9

Callaert, J., Vervenne, J., Van Looy, B., Magermans, T., Song, X., & Jeuris, W. (2014). *Patterns of Science-Technology Linkage*. Retrieved from

Cassiman, B., Veugelers, R., & Zuniga, P. (2008). In search of performance effects of (in)direct industry science links. *Industrial and Corporate Change, 17*(4), 611-646. doi:10.1093/icc/dtn023

Cockburn, I. M., & Henderson, R. M. (1996). Public–private interaction in pharmaceutical research. *Proceedings of the National Academy of Sciences, 93*(23), 12725-12730. Retrieved from http://www.pnas.org/content/93/23/12725.abstract

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations. *Research Policy, 29*(2), 109-123. doi:http://dx.doi.org/10.1016/S0048-7333(99)00055-4

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal, 25*(8-9), 909-928. doi:10.1002/smj.384

Freeman, C. (1987). *Technology, policy, and economic performance : lessons from Japan*. London ; New York: Pinter Publishers.

Gambardella, A. (1992). Competitive advantages from in-house scientific research: The US pharmaceutical industry in the 1980s. *Research Policy, 21*(5), 391-407. doi:https://doi.org/10.1016/0048-7333(92)90001-K

Geuna, A., & Muscio, A. (2009). The Governance of University Knowledge Transfer: A Critical Review of the Literature. *Minerva, 47*(1), 93-114. doi:10.1007/s11024-009-9118-2

Gittelman, M., & Kogut, B. (2003). Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns. *Management Science, 49*(4), 366-382. doi:10.1287/mnsc.49.4.366.14420

Hicks, D., Breitzman, A., Hamilton, K., & Narin, F. (2000). Research excellence and patented innovation. *Science and Public Policy, 27*(5), 310-320. doi:10.3152/147154300781781805

Jaffe, A. B. (1989). Real Effects of Academic Research. *American Economic Review, 79*(5), 957-970. Retrieved from http://www.jstor.org/stable/1831431

Laursen, K., & Salter, A. (2006). Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal, 27*(2), 131-150. doi:doi:10.1002/smj.507

Magerman, T., Van Looy, B., & Song, X. (2010). Exploring the feasibility and accuracy of Latent Semantic Analysis based text mining techniques to detect similarity between patent documents and scientific publications. *Scientometrics, 82*(2), 289-306. doi:10.1007/s11192-009-0046-6

Mansfield, E. (1998). Academic research and industrial innovation: An update of empirical findings. *Research Policy, 26*(7), 773-776. doi:https://doi.org/10.1016/S0048-7333(97)00043-7

Nelson, R. R. (1993). *National innovation systems: a comparative analysis*. New York: Oxford University Press.

Verhoeven, D., Bakker, J., & Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy*, *45*(3), 707-723. doi:http://dx.doi.org/10.1016/j.respol.2015.11.010

Wang, J., Veugelers, R., & Stephan, P. E. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, *46*(8), 1416-1436. doi:https://doi.org/10.1016/j.respol.2017.06.006