NBER WORKING PAPER SERIES

TRACING THE LINKAGES BETWEEN SCIENTIFIC RESEARCH AND ENERGY INNOVATIONS: A COMPARISON OF CLEAN AND DIRTY TECHNOLOGIES

Robert K. Perrons
Adam B. Jaffe
Trinh Le

Working Paper 27777
http://www.nber.org/papers/w27777

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2020

No external funding. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Tracing the Linkages Between Scientific Research and Energy Innovations: A Comparison of Clean and Dirty Technologies
Robert K. Perrons, Adam B. Jaffe, and Trinh Le
NBER Working Paper No. 27777
September 2020
JEL No. O13,O31

ABSTRACT

The challenge of mitigating climate change has focused recent attention on basic scientific research feeding into the development of new energy technologies (Popp, 2017). Energy innovation tends to consist of a series of partially overlapping processes involving: (1) the production of scientific and technological knowledge, (2) the translation of that knowledge into working technologies or artifacts, and (3) the introduction of the artifacts into the marketplace, where they are matched with users' requirements. However, relatively little data are available showing how long each of these processes takes for energy technologies. Here we combine information from patent applications with bibliographic data to shine light on the second process —that is, the translation of scientific knowledge into working prototypes. Our results show that "clean" energy technologies are more dependent on underlying science than "dirty" technologies, and that the average lag between publication of scientific findings and the incorporation of those findings in clean energy patents has risen from about five to about eight years since the 1980s. These findings will help policymakers to devise more effective mechanisms and strategies for accelerating the overall rate of technological change in this domain.

Robert K. Perrons
QUT School of Management
GPO Box 2434
Brisbane, QLD 4001
Australia
robert.perrons@qut.edu.au

Trinh Le
Motu Economic and Public Policy Research
PO Box 24390
Wellington 6142
New Zealand
trinh.le@motu.org.nz

Adam B. Jaffe
188 Brookline Avenue
Apartment 26A
Boston, MA 02215
and Brandeis University and Queensland University of Technology
and also NBER
adam.jaffe@motu.org.nz

1. Introduction

Transitioning the world's energy system away from carbon-intensive sources such as fossil fuels towards non-carbon alternatives represents an immense technological challenge (Deutch, 2011; Henderson & Newell, 2011b; Newell, 2011; Patt, 2015). Despite the urgent case for bringing new energy-related innovations to the marketplace, developing and deploying them on a large scale has been frustratingly slow (Bunn, Anadon, & Narayanamurti, 2014; Grubler, 2012), often taking several decades to develop even the most promising ideas into novel technologies that achieve a significant amount of market penetration (Lester & Hart, 2012). But even with the broad consensus about the protracted timeframes required for these activities (Smil, 2010, 2017), there is little evidence on the source or nature of the delays. Significant attention has been paid to where and how financial investments in clean energy technologies occur (Lerner, 2011; Nanda, Younge, & Fleming, 2015; Rotman, 2019; Yeo, 2019) but, aside from a handful of case studies focused on individual technologies (e.g., Grubler & Wilson, 2014), far less is understood about the investments of time that are required to develop and deploy these innovations. In this paper, we build upon Popp's (2017) examination of the flows of knowledge between universities, the private sector, and government agencies on the way to the development of new energy technologies.

However, whereas Popp (2017) mostly relied on a manual analysis of the data, this paper uses the Lens, a new online tool that links information from patent applications with bibliographic data, to usefully extend Popp's earlier findings. Specifically, we present evidence showing how much time passes in the energy domain between the production of scientific knowledge and the delivery of working technologies or artifacts. We start by reviewing prior investigations in this area, explaining the role of patents in technological innovation, and putting forward the research questions at the center of this investigation. We then present evidence quantifying the amount of time required to translate new science into energy-related innovations, and conclude with a brief discussion about the importance and policy-related implications of these findings.

2. Background and Research Questions

The long time horizons required for energy innovation are frequently discussed in the literature in a vague and aggregated way, but innovation tends to consist of a series of partially overlapping processes (Pavitt, 2005; Wilson & Grubler, 2014). As Figure 1 shows, technological innovation typically begins with the production of scientific and technological knowledge, which is then translated into working technologies or artifacts. The artifacts typically require further development to be viable in the marketplace, including adaptation to users' needs and scale-up from laboratory prototypes to viable volume. There is, of course, tremendous attrition across these processes, as much new knowledge never leads to any working artifacts, and many inventions that are manifest in working artifacts are never transitioned to a market innovation. But from the perspective of a successful market innovation, we can typically look backward and see the prototype invention from which it developed, and the underlying scientific and technological knowledge that contributed to that invention.

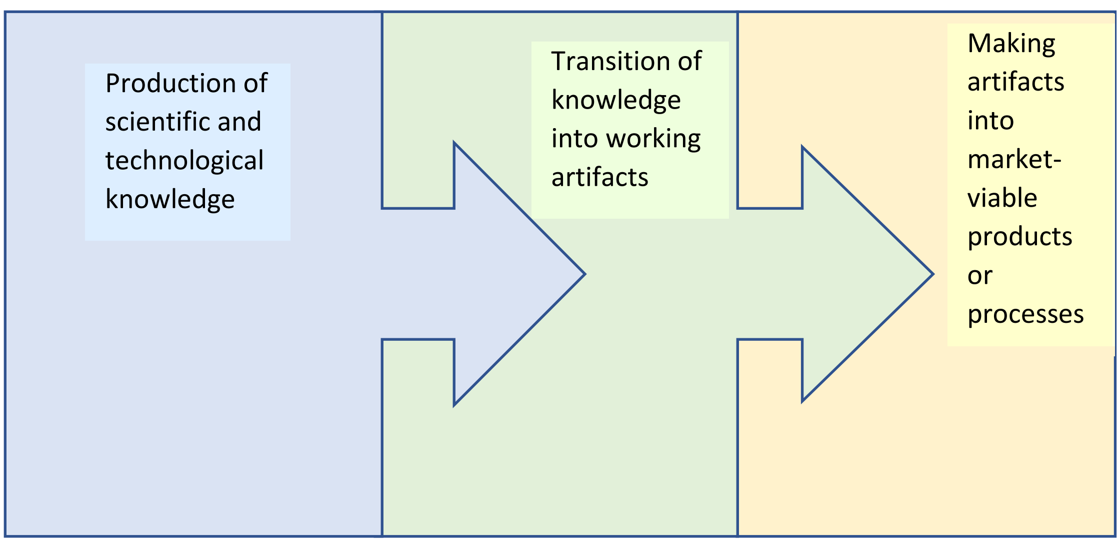

Figure 1: Processes of Innovation (based on Pavitt (2005) pp. 88-89)

Each of these processes involves a different combination of stakeholders who respond to different incentives and face different constraints. As a result, different policy instruments and strategies have been recommended for each in an attempt to accelerate the overall clean energy technology innovation journey (Deutch & Lester, 2004; Grubler & Wilson, 2014; Henderson & Newell, 2011a). It is therefore useful to pry apart the overall timeframe required for energy innovation into its constituent parts so that the length of each process can be examined separately. In this Research Note, we zoom in on the middle process in Figure 1, leading to:

> **Research Question 1**: How important is basic scientific research as a foundation for new clean energy technology?
>
> **Research Question 2**: What is the magnitude of the lags between the production of scientific knowledge and the delivery of clean energy working technologies or artifacts?

3. Methodology

3.1 The Role of Non-Patent Literature in Patent Applications

The patent system provides a useful window into the second stage of Figure 1, and the connections between the first stage and the second stage. A patent represents a legal acknowledgement that an invention has occurred. The patent document captures key aspects of this invention, and places that invention in the context of relevant knowledge at the time a patent is filed (Jefferson et al., 2018). Applicants find and disclose evidence of science and technology that preceded the work and upon which the innovation is based, and the patent examination process uncovers additional "prior art." Much of this disclosure is in the form of previous patents, but it also includes non-patent literature (NPL) such as publications in scholarly journals. NPL is an indicator of the feedstock science that underpins the technological concept underlying the patent application (Jefferson et al., 2018). Therefore, patent citations to NPL provide traces of pathways connecting the first and second processes described above.

Following earlier categorizations of energy technology patents (Bosetti & Verdolini, 2013; Lanzi, Verdolini, & Haščič, 2011), we divided all the international patent classifications (IPCs) related to energy production into two groups: "dirty" and "clean." For each NPL item identified in the prior art search for each of these patents, we then measured the amount of time that had passed between when the scientific literature referenced in an energy-related patent application was published, and when the resulting patent application had been filed.[2]

[2] Citations here only include scholarly works for which a DOI, PMID, Microsoft Academic ID, or PMCID can be found. The analysis therefore excludes other kinds of publications such as working papers, brochures, works published by outlets which are not covered by the above sources, or works that cannot be found in the above sources due to insufficient or incorrect information provided in the patent.

We recognize that this approach affords only an incomplete and imperfect representation of the contributions of scientific and technological knowledge to the generation of "working artifacts." First, the prior art search that yields the NPL references serves a specific legal function, and does not necessarily identify the broad scope of knowledge inputs of the invention. Second, while patent law requires that an invention be "reduced to practice" in order for a patent to be granted, patents in fact vary greatly in the extent to which a working artifact has truly been achieved. Third, there is frequently only a tenuous connection between the advent of novel technologies and any patents that are ostensibly underlying them (Jaffe & Lerner, 2004). Finally, many inventions are never patented at all, so there will be no traces of their development in the patent data. Despite these limitations, we believe that that the NPL citations in energy patents provide a useful window into the scale and timing of linkages between scientific research and new technologies in the energy domain. A large fraction of the participants and contributors connected to these processes do use patents as a way to defend their intellectual property rights (Henderson & Newell, 2011a), so the picture presented by these data is an important indicator of these processes despite being incomplete.

Patent references to NPL have been used previously to study how knowledge flows between and among universities, the private sector, and government research institutions contribute to energy innovation (Popp, 2017). We build upon this earlier contribution by examining both the extent of citation to NPL and the lag between NPL references and patent applications reflected in "clean" as compared to "dirty" energy patents. The data were extracted from the Lens (www.lens.org), an online tool that connects patent databases around the world with scholarly and technical literature. Through it, we were able to access a broad range of information about worldwide patent applications and the scientific literature that underpins them.

3.2 Categorizing Patents

The patents used in this study were classified according to the IPC system of the World Intellectual Property Organization (WIPO).[3] Prior investigations in the area of energy innovations (e.g., Ardito, Petruzzelli, & Albino, 2016; Bosetti & Verdolini, 2013; Johnstone, Haščič, & Popp, 2010; Lanzi et al., 2011) have not entirely agreed on what IPC codes in the energy domain are "clean" and which ones are "dirty."

---

[3] https://www.wipo.int/classifications/ipc/en - accessed on 6 February 2020.

In this investigation, we adopted Bosetti et al.'s (2013) list of IPC codes for clean technologies, and included in our dirty technologies category IPC codes that appeared in either Bosetti et al.'s (2013) or Lanzi et al.'s (2011) lists of dirty technologies. We framed our classifications in this way because the above categorizations offer an established classification system of IPC codes that focus on energy production technologies. Several of the other classification systems were inclusive of patent groups like energy conservation or nuclear power generation for which there were no corresponding "dirty" innovations. By confining our analysis only to IPC codes related to energy production technologies, we were able to make reasonable like-for-like comparisons between the clean and dirty categories.

3.3 Linking IPCs to NPL

The IPCs were then submitted to the Lens[4], an online tool that connects patent databases around the world with scholarly and technical literature, to identify all of the individual patents applications filed during the period 1970-2015.[5] The data revealed that about 2.8% of the patents within the study have any form of NPL connected to them that can be found on the Lens scholarly database. Over 99% of these cited works had at least one digital object identifier (DOI), while the remaining 1% had a PubMed identification (PMID), Pubmed Central identifier (PMCID), or Microsoft Academic ID.

3.4 Reference Data

Within the Lens, we limited the query to "citing patents," and then chose the "cited works" option to show the metadata of the cited scholarly works connected to each of the patent applications within the targeted IPC codes. The following data fields were provided for each cited work:

a) Article title
b) Publication year
c) Type of document
d) Name of journal
e) Volume and issue number of journal

[4] www.lens.org

[5] By using the Lens' Boolean search function, we were able to query multiple IPC codes at a time. One limitation of the Lens at the time of this investigation, however, was that users were only able to export results that had no more than 50,000 records per query. The 207 IPC codes for the "clean" and "dirty" categories resulted in more than 50,000 records, so we needed to break apart the data gathering into seven separate queries, and then consolidate them into a joined-up data set for further analysis.

f) Publisher
g) Names of authors
h) Country of publisher

The resulting database included bibliographic information for over 18,000 different NPLs. We conducted two-sided t-tests to determine whether there was a difference between clean and dirty energy technologies with regards to (1) the average number of NPL citations per patent application, and (2) the time lags between patent applications and the NPL citations within them. The results for these t-tests are shown in the Appendix, and are also depicted graphically in Figures 2 and 3.

4. Results and Discussion

In Figures 2-4, the bar height represents the mean (fraction of patents, number of papers or time lag) and the error bars represent the 95% confidence intervals. Figure 2 shows the fraction of patents (clean and dirty) that contain any reference to NPL, by decade. Figure 3 shows, for those patents that do have NPL references, the mean number of such NPL items referenced in each awarded patent. Taking the number of NPL per patent as an indicator of the importance of science as an input to invention, Figures 2 and 3 show two important things. First, the frequency of NPL citations in energy patents has been increasing on both the extensive and intensive margins, i.e. a larger share of patents cite NPL and those that do cite NPL have a greater number of NPL citations per patent. It is likely that this trend is, to some extent, an artifact of technological and organizational changes that make it easier for patent examiners to find non-patent prior art. But it also likely that increases in the extent of NPL citation indicate an increasing dependence of invention on underlying scientific research over time (Jaffe and De Rassenfosse, 2019). That is, energy invention (both clean and dirty) has become increasingly dependent on science in the last two decades relative to the end of the 20th century.

The other clear pattern in Figures 2 and 3 is that the dependence of clean energy technologies on science is greater than that of dirty technologies, and again this is true on both the extensive and intensive margins. The fraction of patents with any science connection—as indicated by a NPL reference—is consistently approximately twice as high for clean patents as for dirty patents. And, for those patents in either group that do cite NPL, the NPL references per patent is greater for clean technologies in all periods. The period-by-period difference is statistically significant only for the 2000s but, if the data for all periods is combined, the difference is statistically significant. (See Appendix for t-tests on these differences.) The relatively high and rising importance of science for clean energy invention reinforces the urgency of investment in the underlying basic science as a crucial component of global strategy to manage climate change (Jaffe, 2012; National Research Council, 2010).

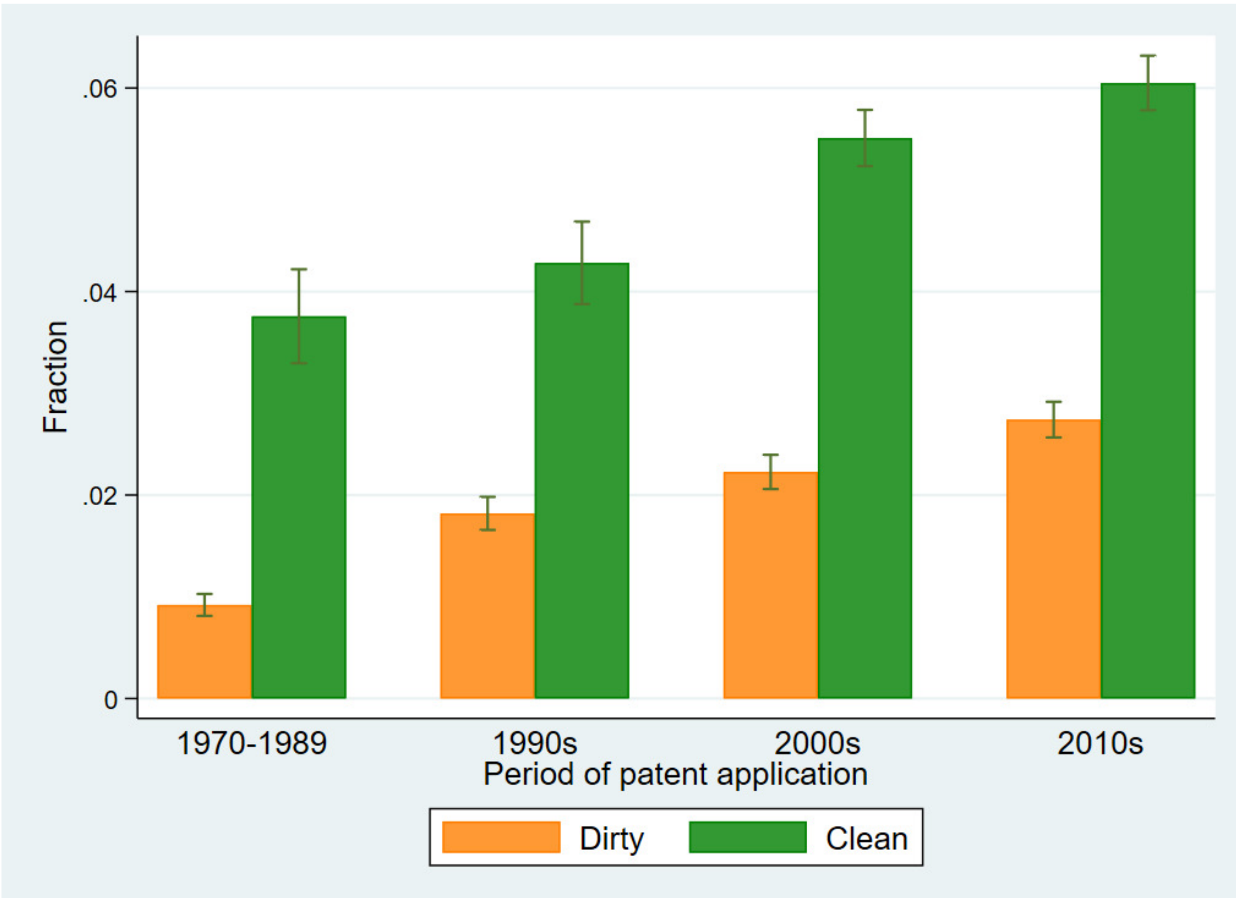

Figure 2: Fraction of Patents Citing NPL by Period

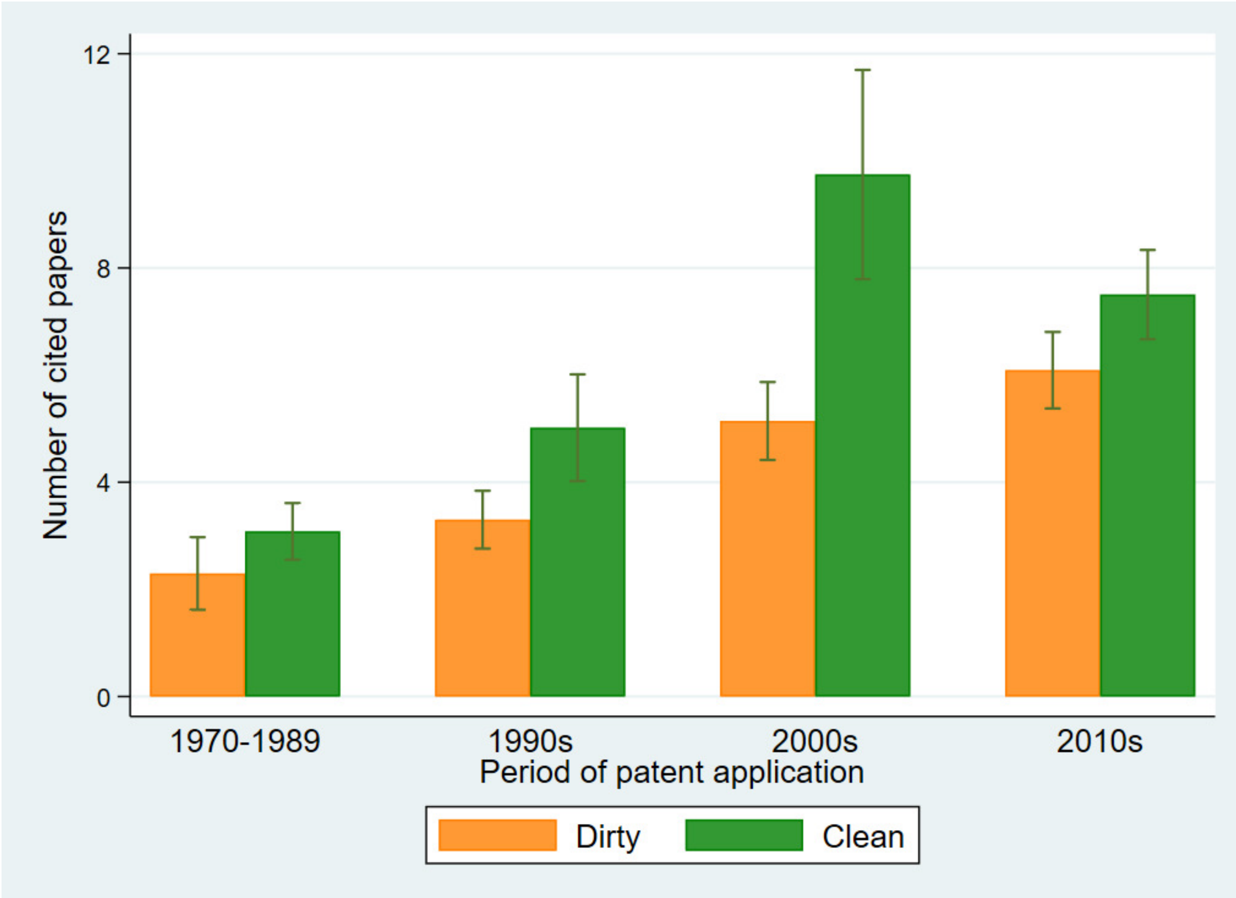

Figure 3: Average Number of Papers Cited per Patent by Period

Figure 4 shows the average time lag between a published paper and a patent containing a NPL reference to that paper. The "time lag" variable is the number of years between when the cited scholarly article was published and when the patent was applied for. However, a patent might cite several articles, and similarly an article might be cited by several patents. Two weighting methods were applied to consolidate multiple linkages between patent applications and NPLs into a single calculation. In the "plutocratic" weighting method, each patent-paper combination is weighted equally. That is, if a patent cites N papers, then it will be used N times in the calculations (i.e., once for each cited paper). In the "democratic" weighting method, if a patent cites N papers, then each patent-paper combination for that patent only carries a weight of 1/N, thereby ensuring that the total weight for each patent is always 1. The plutocratic method gives more weight to newer patents than does the democratic method, as newer patents tend to cite more than older ones. The results presented in Figure 4 are based on the plutocratic weighting method.

Figure 4 shows that, during the past decade, it took on average approximately 8 years for scientific discoveries to appear in patent applications related to clean energy technologies. Interestingly, the average for dirty energy technologies in the most recent decade was about 11 years. This may reflect a tendency for clean technologies to rely somewhat more on the most recent science, while dirty technologies have more of a reliance on older science. But whatever the factors underlying the difference, these relatively long lags suggest that the time needed for the middle stage of Figure 1 represents a significant fraction of the overall time lags necessary to bring new cleaner technologies into actual use.

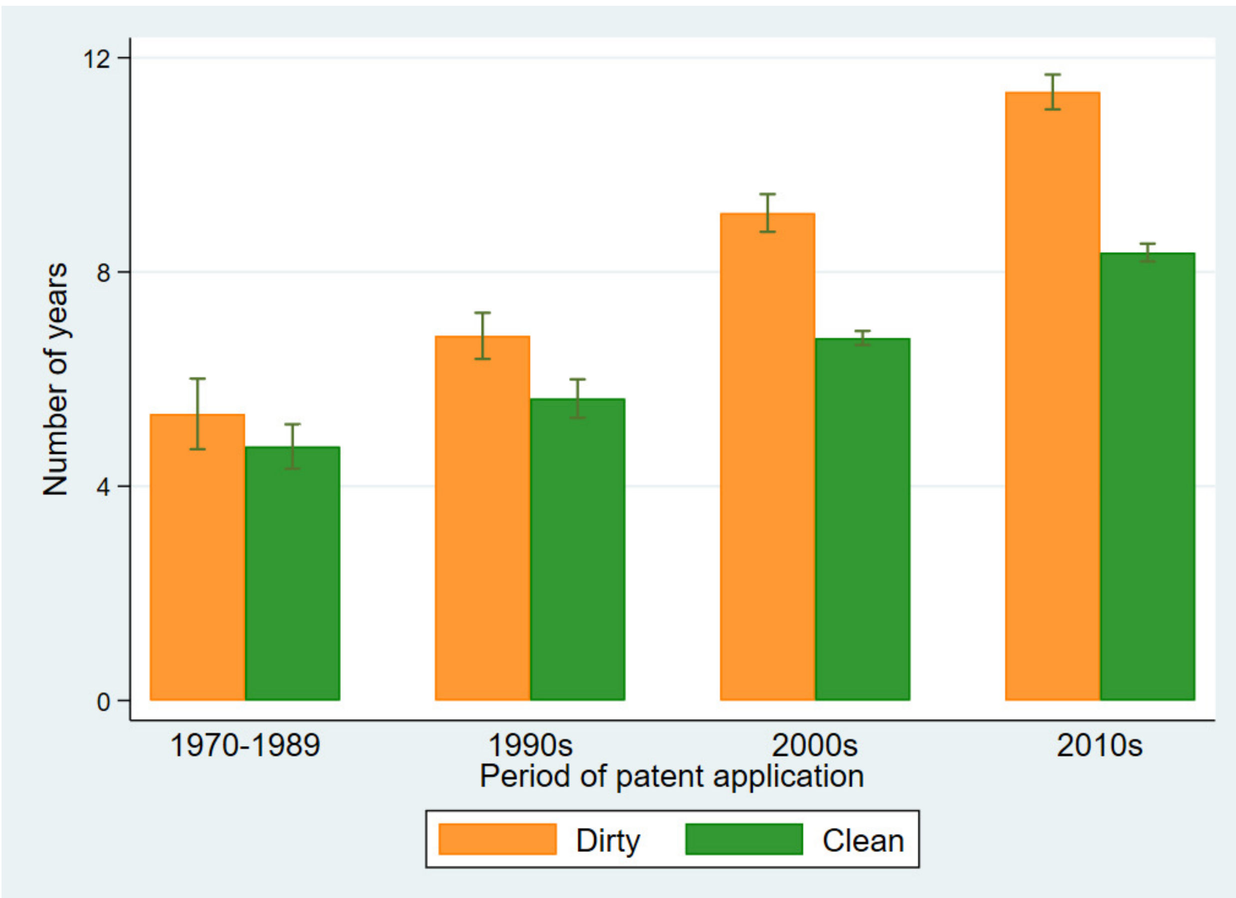

Figure 4: Years between Academic Paper and Patent Application

5. Conclusions and Policy Implications

This finding carries with it important consequences for policymakers: barring any unexpected departures from this pattern, discoveries arising from investments in science that are made today to help the transition towards clean energy will quite likely take many years to find their way into new energy technologies. Conventional wisdom has traditionally held that the capital-intensive nature of the energy sector is a significant contributing factor to how slowly it is able to develop and deploy technological innovations(Lester & Hart, 2012), but our evidence shows that a significant part of the problem lies further upstream before the capital-intensive nature of the industry is materially relevant. This finding also highlights the time-critical nature of investments in clean energy technologies. If significantly new technologies will be required to mitigate climate change in a useful timeframe (Jaffe, 2012), then the scientific research laying the foundations for these new technologies needs to be undertaken very soon.

Figure 4 also reveals that the process of turning scientific and technological knowledge into working artifacts is not getting shorter, despite the chorus of voices that has been calling over many years for the acceleration of energy-related innovation (Henderson & Newell, 2011a). It should be emphasized, too, that this figure only considers the middle process of the three shown in Figure 1, and includes neither the amount of time it will take for these investments in science to actually generate published results (i.e., the first process), nor how long it will take for the new technology to be matched with user requirements in the marketplace (i.e. the third process).

As energy technologies become more science-intensive, it follows that the success of processes for creating new energy technologies will become even more dependent on how scientific knowledge is created and integrated into these innovations. By improving the granularity of our understanding of these processes, the evidence presented here will help policymakers and to make more effective mechanisms and strategies for accelerating the overall rate of technological change in this domain.

6. Acknowledgments

For help with preliminary data used in a previous draft, we thank Mr. Sander Fels-Leung for writing the Java program used to acquire the data from the Lens.org and compile the resulting data into Excel

spreadsheets, and Professor Gaétan de Rassenfosse and Dr. Orion Penner for their expert advice in writing the Java code in a way that successfully connected with the digital architecture of the Lens.org.

7. Author Contributions

RKP and ABJ put together the conceptual approach for the paper and managed the preliminary stages of the data collection process from the Lens. TL updated the data, analyzed it, summarized the findings, and assessed the robustness of the results. RKP led the writing of the text, and ABJ and TL made significant additions and edits.

8. Data and Code Availability

All data analyzed during this study are available from authors on request.

9. References

Ardito, L., Petruzzelli, A. M., & Albino, V. (2016). Investigating the Antecedents of General Purpose Technologies: A Patent Perspective in the Green Energy Field. *Journal of Engineering and Technology Management, 39*, 81-100.

Bosetti, V., & Verdolini, E. (2013). *Clean and Dirty International Technology Diffusion*. Fondazione Eni Enrico Mattei working paper no. 43.2013. Milan, Italy.

Bunn, M., Anadon, L. D., & Narayanamurti, V. (2014). The Need to Transform U.S. Energy Innovation. In L. D. Anadon, M. Bunn, & V. Narayanamurti (Eds.), *Transforming U.S. Energy Innovation* (pp. 1-35). New York: Cambridge University Press.

Deutch, J. M. (2011). *The Crisis in Energy Policy*. Cambridge, Massachusetts: Harvard University Press.

Deutch, J. M., & Lester, R. K. (2004). *Making Technology Work: Applications in Energy and the Environment*. New York: Cambridge University Press.

Grubler, A. (2012). Energy Transitions Research: Insights and Cautionary Tales. *Energy Policy, 50*, 8-16.

Grubler, A., & Wilson, C. (2014). *Energy Technology Innovation*. New York: Cambridge University Press.

Henderson, R. M., & Newell, R. G. (2011a). *Accelerating Energy Innovation: Insights from Multiple Sectors*. Chicago: University of Chicago Press.

Henderson, R. M., & Newell, R. G. (2011b). Introduction and Summary. In R. M. Henderson & R. G. Newell (Eds.), *Accelerating Energy Innovation: Insights from Multiple Sectors* (pp. 1-23). Chicago: University of Chicago Press.

Jaffe, A. B. (2012). Technology Policy and Climate Change. *Climate Change Economics, 3*(4), 1250025.

Jaffe, A. B., & De Rassenfosse, G. (2019). "Patent citation data in social science research: Overview and best practices." Research Handbook on the Economics of Intellectual Property Law. Edward Elgar Publishing.

Jaffe, A. B., & Lerner, J. (2004). *Innovation and Its Discontents*. Princeton, New Jersey: Princeton University Press.

Jefferson, O. A., Jaffe, A., Ashton, D., Warren, B., Koellhofer, D., Dulleck, U., . . . Jefferson, R. A. (2018). Mapping the Global Influence of Published Research on Industry and Innovation. *Nature Biotechnology, 36*(1), 31-39.

Johnstone, N., Haščič, I., & Popp, D. (2010). Renewable Energy Policies and Technological Innovation: Evidence Based on Patent Counts. *Environmental and Resource Economics, 45*(1), 133-155.

Lanzi, E., Verdolini, E., & Haščič, I. (2011). Efficiency-Improving Fossil Fuel Technologies for Electricity Generation: Data Selection and Trends. *Energy Policy, 39*(11), 7000-7014.

Lerner, J. (2011). Venture Capital and Innovation in Energy. In R. M. Henderson & R. G. Newell (Eds.), *Accelerating Energy Innovation: Insights from Multiple Sectors* (pp. 225-260). Chicago: University of Chicago Press.

Lester, R. K., & Hart, D. M. (2012). *Unlocking Energy Innovation: How America Can Build a Low-Cost, Low-Carbon Energy System*. Cambridge, Massachusetts: MIT Press.

Nanda, R., Younge, K., & Fleming, L. (2015). Innovation and Entrepreneurship in Renewable Energy. In A. B. Jaffe & B. F. Jones (Eds.), *The Changing Frontier: Rethinking Science and Innovation Policy* (pp. 199-232). Chicago: University of Chicago Press.

National Research Council. (2010). *Limiting the Magnitude of Future Climate Change*. Retrieved from Washington, DC:

Newell, R. G. (2011). The Energy Innovation System: A Historical Perspective. In R. M. Henderson & R. G. Newell (Eds.), *Accelerating Energy Innovation: Insights from Multiple Sectors* (pp. 25-47). Chicago: University of Chicago Press.

Patt, A. (2015). *Transforming Energy: Solving Climate Change with Technology Policy*. New York: Cambridge University Press.

Pavitt, K. (2005). Chapter 4: Innovation Processes. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), *Oxford Handbook of Innovation* (pp. 86-114). New York: Oxford University Press.

Popp, D. (2017). From Science to Technology: The Value of Knowledge from Different Energy Research Institutions. *Research Policy, 46*(9), 1580-1594.

Rotman, D. (2019). What Would You Pay to Save the World? *MIT Technology Review, 122*(3), 8-11.

Smil, V. (2010). *Energy Myths and Realities: Bringing Science to the Energy Policy Debate*. Washington, DC: American Enterprise Institute.

Smil, V. (2017). *Energy Transitions: Global and National Perspectives* (2nd ed.). Santa Barbara, California: Praeger.

Wilson, C., & Grubler, A. (2014). Energy Technology Innovation. In A. Grubler & C. Wilson (Eds.), *Energy Technology Innovation* (pp. 3-10). New York: Cambridge University Press.

Yeo, S. (2019). Climate Finance: The Money Trail. *Nature, 573*, 328-331.

10. Appendix – Statistical Significance of Results

Period	Number of Applications for "Dirty" Patents (whether citing papers or not)	Fraction of "Dirty" Patents Citing NPL	Number of Applications for "Clean" Patents (whether citing papers or not)	Fraction of "Clean" Patents Citing NPL	t-statistic on the difference between clean and dirty
1970-1989	30117	0.01	6521	0.04	-17.602
1990s	25999	0.02	9599	0.04	-13.281
2000s	29593	0.02	26140	0.06	-20.385
2010s	33273	0.03	30315	0.06	-20.569
Total	118982	0.02	72575	0.05	-20.569

Table A1: Results and Statistical Significance of Data Behind Figure 2

Period	Number of Applications for "Dirty" Patents that cite NPL	Average Number of Papers Cited by "Dirty" Patents (conditional on citing)	Number of Applications for "Clean" Patents that cite NPL	Average Number of Papers Cited by "Clean" Patents (conditional on citing)	t-statistic on the difference between clean and dirty
1970-1989	277	2.30	245	3.08	-1.768
1990s	473	3.30	411	5.02	-3.083
2000s	659	5.14	1440	9.74	-3.082
2010s	912	6.09	1834	7.50	-2.170
Total	2321	4.76	3930	7.80	-5.289

Table A2: Results and Statistical Significance of Data Behind Figure 3

Period	Number of Applications for "Dirty" Patents that cite NPL	Average Number of Years between Paper and "Dirty" Patent Application	Number of Applications for "Clean" Patents that cite NPL	Average Number of Years between Paper and "Clean" Patent Application	t-statistic on the difference between clean and dirty
1970-1989	277	5.35	245	4.74	1.583
1990s	473	6.81	411	5.64	4.116
2000s	659	9.10	1440	6.77	14.443
2010s	912	11.36	1834	8.36	17.662
Total	2321	9.62	3930	7.31	21.876

Table A3: Results and Statistical Significance of Data Behind Figure 4