NBER WORKING PAPER SERIES

FROM SCIENCE TO TECHNOLOGY: THE VALUE OF KNOWLEDGE FROM DIFFERENT ENERGY RESEARCH INSTITUTIONS

David Popp

Working Paper 22573
http://www.nber.org/papers/w22573

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2016

This work was supported by NSF grant # SMA-1064161. Jung Eun Kim, Qing Miao and Tian Tang all provided excellent research assistance on this project. I thank Mort Webster for comments on earlier versions of this research. Finally, I thank Ann Beynon and Lauren Cygler of Thomson Reuters for their assistance in obtaining the publication data used in this paper. Certain data included herein are derived from the Web of Science Core Collection ® prepared by THOMSON REUTERS ®, Inc. (Thomson ®), Philadelphia, Pennsylvania, USA: © Copyright THOMSON REUTERS ® 2012. All rights reserved. Thomson Reuters reserves the right to review the paper prior to dissemination to ensure that the confidentiality of the data is not unintentionally compromised. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

From Science to Technology: The Value of Knowledge From Different Energy Research Institutions
David Popp
NBER Working Paper No. 22573
August 2016
JEL No. O38,Q42,Q48,Q55

ABSTRACT

Using an original data set of both scientific articles and patents pertaining to alternative energy technologies, this paper provides new evidence on the flows of knowledge between university, private sector, and government research. Better understanding of the value of knowledge from these institutions can help decision makers target R&D funds where they are most likely to be successful. I use citation data from both scientific articles and patents to answer two questions. First, what information is most useful to the development of new technology? Does high quality science lead to commercial success? I find that this is the case, as those articles most highly cited by other scientific articles are also more likely to be cited by future patents. Second, which institutions produce the most valuable research? Are there differences across technologies? Research performed at government institutions appears to play an important translational role linking basic and applied research, as government articles are more likely to be cited by patents than any other institution, including universities. Universities play a less important role in wind research than for solar and biofuels, suggesting that wind energy research is at a more applied stage where commercialization and final product development is more important than basic research.

David Popp
Department of Public Administration
and International Affairs
Syracuse University
The Maxwell School
426 Eggers Hall
Syracuse, NY 13244-1020
and NBER
dcpopp@maxwell.syr.edu

1. Introduction

Developing new and improved clean-energy technologies is an important part of any strategy to combat global climate change. For example, generation of electricity and heat is the largest source of carbon emissions, accounting for 42 percent of carbon emissions worldwide in 2012 (IEA 2014). Meeting the climate policy goals currently under consideration, such as European Union discussions to reduce emissions by 40 percent below 1990 levels by 2030 or the U.S. Clean Power Plan goal of reducing emissions from the electricity sector by 32 percent by 2030, will not be possible without replacing much of the current fossil fuels-based electric generating capacity with alternative, carbon-free energy sources.

Because clean energy technologies are usually not competitive with fossil fuels without policy support (Greenstone and Loney, 2012), a large academic literature has emerged evaluating the role of environmental policy for fostering clean energy innovation. Much of this research focuses on the private sector, showing that both higher energy prices and targeted support for renewable energy, such as feed-in tariffs or renewable portfolio standards, lead to increases in clean energy patents.[1]

Even when environmental regulations encourage private sector innovation, firms will focus research efforts on technologies that are closest to market (Johnstone *et al*. 2010). Yet, one challenge facing many climate-friendly innovations is the long time-frame from the initial invention to successful market deployment. Consider, for instance, the case of solar energy. Despite research efforts that began during the energy crises of the 1970s, solar is still only cost competitive without policy support in niche markets, such as remote off-grid locations. This leaves a role for government-sponsored R&D to fill in the gaps, particularly in the case of

[1] Examples include Johnstone et al (2010), Verdolini and Gaelotti (2011), Peters *et al.* (2012), Veugelers (2012), Dechezleprêtre and Glachant (2014), Nesta et al. (2014). Dechezleprêtre and Popp (2014) and Popp (2011) provide recent reviews.

climate change, where a diversified energy portfolio will be necessary to meet currently proposed emission reduction targets. Recognizing this need, during the December 2015 Paris climate meetings, a coalition of governments pledged to double their renewable energy R&D budgets to $20 billion over the next five years (Davenport and Wingfield, 2015).

While many studies have looked at private sector energy R&D, fewer papers address the effectiveness of public sector clean energy R&D. Those that do typically find a positive effect of public R&D on patenting (e.g. Johnstone et al 2010, Verdolini and Gaelotti 2011, Dechezleprêtre and Glachant 2014, Nesta et al. 2014). However, these studies typically include just a single lagged value of energy R&D, raising questions about what is truly identified.[2]

To better ascertain the effectiveness of public energy research, Popp (2016) links data on scientific publications to public energy R&D funding. For evaluating public research funding efforts, publication data provide a more appropriate outcome measure than patents. By looking at the effect of public R&D funding on scientific articles, Popp (2016) isolates the effect of public R&D to shed light on the process through which public R&D helps develop scientific knowledge. As the ultimate goal of government energy R&D funding is not an article, but rather a new technology, Popp uses citations to link these articles to new energy patents. While public funding does lead to new articles, lags in both the creation of a new publication and the transfer of this knowledge to applied work mean that public R&D spending takes several years to go from new article to new patent.

While Popp (2016) focuses on the time it takes for the results of public R&D to be cited by a new patent, this paper extends that work by providing more detail on the knowledge flows between published and patented clean energy research. Given recent calls for more public

[2] A partial exception is Peters *et al*. (2012), who state that they test multiple lags and stocks of public R&D in unreported results.

energy R&D efforts, such as the aforementioned pledges at the December 2015 Paris climate meetings, identifying the investments most valuable to further advancing energy research can help decision makers target R&D funds towards both the technologies and institutions where they are most likely to be successful. This paper uses citation data from both scientific articles and patents to answer two questions:

1) What information is most useful to the development of new technology? Does high quality science lead to commercial success? That is, are scientific articles cited frequently by other articles also more likely to be cited by patents, or are the types of articles cited by patents different from those cited by articles? Popp (2016) argues that there is room to expand public R&D budgets, as there is little change in either the quantity of published research or the quality, as measured by citations, after large increases in public energy R&D. But, are citations within the published literature an appropriate measure of the relevance of this published research for applied work? In Section 3, I show that highly cited journal articles do receive more citations by patents, suggesting that counts of journal-to-journal citations are a good indicator of the ultimate value of an article for technology development.

2) Which institutions produce the most valuable research? Are there differences across technologies? Using patent and article citations as a measure of knowledge flows, in section 4 I ask which institutions provide the most useful building blocks for future researchers. Do collaborations between public and private research organizations increase flows of knowledge among groups? As governments prepare to expand renewable energy R&D, such evidence can inform where public research funds can best be targeted. While government research efforts are often criticized as wasteful, I find that government patents are cited more frequently by researchers than other patents, and that government research articles are more likely to be cited

by future patents. Thus, government research does appear to play an important translational role linking basic and applied research. Universities play a less important role in wind research than for solar and biofuels, suggesting that wind energy research is at a more applied stage where commercialization and final product development is more important than basic research.

2. Data

In this paper, I use scientific articles to represent trends in more basic upstream research and I use patents to represent trends in more applied downstream research and product development. The article data come from the Thomson Reuters Web of Science Core Collection database. Using a series of keyword searches of article titles, abstracts, and keywords, provided in Appendix Table A1, I identified journal publications for each of three technologies: biofuels, solar energy, and wind energy.[3] I focus on publications in scientific journals by dropping articles such as reviews, editorials, or news items. I do include proceedings papers that are included in the Web of Science Core Collection database. The article data run from 1991-2011, as complete records of titles, abstracts, and keywords are first available in 1991. Having identified appropriate keywords, Thomson Reuters provided a custom database containing all articles from 1991-2011 for each technology. The database includes descriptive information on each paper, including the date of publication and addresses for each author, which I use to assign articles to each country. For each technology, articles are aggregated by year of publication and country. In the case of articles with multiple authors from multiple countries, I use weighted counts, assigning articles proportionately by the number of countries represented. For example, an article with 2 US authors and 1 Japanese author counts as 0.67 for the US and 0.33 for Japan.

[3] The choice of technologies follows from Popp (2016). There, a key factor limiting relevant technologies was the availability of a sufficient time series of public R&D spending on each technology. In addition, given concerns about climate change, the research focuses on alternative energy technologies rather than fossil fuels.

In addition, the database also includes descriptive data on each article citing these energy articles.

When developing keyword searches, there is a tradeoff between using broad searches that identify as many relevant articles as possible but also include some irrelevant articles or using narrower searches that filter out irrelevant articles but may miss some relevant ones. I devised searches that would be narrow, so as to avoid irrelevant articles. As such, my database does not necessarily include, for example, every article related to wind energy published since 1991. However, as long as there is no change in the share of relevant articles identified over time, my results will still be an unbiased indicator of the effect of trends in alternative energy research. This simply requires assuming that the searches consistently identify a fixed percentage of wind articles published in any given year. In contrast, using broader search terms that identified more wind articles but also included irrelevant articles would require assuming that the irrelevant articles responded in the same way as actual wind publications included in the sample.

As the ultimate goal of clean energy research is a new technology, I link my publication data to patent data, which reflect the output of applied research efforts. Patents contain citations to scientific articles, allowing direct linkages to be made. Moreover, recent work by Roach and Cohen (2013) shows that references to non-patent literature (NPL) such as journal articles are better measures of knowledge flows from public research to patents than are citations to other patents. Sorenson and Fleming (2004) compare forward patent citations received by patents that cite or do not cite published materials. Those patents citing previously published non-patent literature receive more citations and are cited more quickly than other patents. Thus, focusing on the flow of knowledge between articles and patents highlights research contributing to the most influential patents.

The major challenge for this analysis is linking the article data to patents. Due to data constraints, the patent citation analysis focuses on citations made by U.S. patents. Using the International Patent Classification (IPC) system to identify patents pertaining to specific technologies, I identify patents related to biofuels, solar energy, and wind. Appendix Table A2 presents the list of IPC technologies used. Data on relevant patents come from the on-line database provided by Delphion (http://www.delphion.com/). I obtained both patent and NPL references for these patents, identifying those patents referencing journal articles. As there is no standard form for citing articles in a patent, matching articles and patents was done manually. To be consistent with the article data, I track patenting trends over time using the grant year of the patents.

[FIGURE 1 ABOUT HERE]

[FIGURE 2 ABOUT HERE]

Figures 1 and 2 show the overall trends in patents and scientific articles, respectively. Both have increased over time, due in part to increased public energy R&D funding (Popp 2016) and new renewable energy policy mandates (Johnstone *et al.*, 2010). For biofuels, counts of articles have grown faster than patents, while the opposite is true for wind. Appendix Figures A1-A6 present trends for the top five countries in each technology. The increase in solar energy patents during the early 2000s is due to an increase in solar energy patents in Japan, which is second to only the US for total solar energy patents. Table 1 shows the ten countries with the most patents and articles for each technology. For patents, the top ten is dominated by high income countries, with just a few exceptions (India for biofuels, albeit with just 7.25 patents, and Taiwan for solar and wind). In contrast, emerging countries appear more frequently on the list of top scientific article sources. China is second to only the US in both biofuels and solar energy articles, and sixth for wind. India is third for biofuels and sixth for solar. Brazil is fourth for biofuels. Thus, these emerging economies are performing research on renewable energy, but this research appears less likely to translate into patentable end use technologies, at least for innovations of sufficient value to be patented in the United States.

[TABLE 1 ABOUT HERE]

[TABLE 2 ABOUT HERE]

Table 2 provides basic summary statistics for the number of citations received by each patent or article. The top section provides data on article-to-article citations, the middle section on patent-to-patent citations, and the bottom on patent-to-article citations. Note that there are many more articles than patents for each of the technologies. Thus, not surprisingly, the average number of article-to-article citations is also largest. Interestingly, while there are twice as many solar energy articles as biofuels articles in my sample, the average number of citations received is similar. Note as well that citations are highly skewed. For both article and patent citations, the mean citations is larger than the median, and is comparable for biofuels and solar to the 75th percentile. Many patents receive no citations, with the median being 0 for biofuels, and just one for solar and wind.

Articles receiving citations from renewable energy patents occurs much less frequently. Thus, in the bottom section, I provide descriptive statistics for both all articles and for the subset of articles that do receive at least one patent citation. Just 0.66% of all biofuels articles, 1.31% of solar articles, and 1.14% of wind articles receive even one patent citation. Moreover, conditional on receiving a citation, the median article receives just one patent citation. The average number of citations received is two or less.[4]

3. Do highly cited articles generate applied technology?

Does high quality science lead to commercial success? While researchers commonly used citations as a proxy for article quality, are scientific articles cited frequently by other articles also more likely to be cited by patents? Or, are the articles that inventors of new technology find useful different from those that other academic researchers find useful? For example, might there be "intermediary" publications that link the results of basic science to applied technologies but are not highly cited by other journal articles? To address these questions, in this section I ask whether highly cited scientific articles (e.g. those more likely to be cited by other articles) are also more likely to be cited by future patents.

As many articles have had only a few years to be cited, truncation of the data is a concern. Because of these truncation issues, and because most articles are only cited by a patent once, I use a hazard model to focus on the probability of an article ever receiving a patent citation. To allow time for articles to be cited, I only consider articles published in 2009 or earlier. The patent data extends through 2011. The main variable of interest is the total number of citations each potentially cited article received through 2011 from other journal articles. This variable tests whether the same articles are cited by both journals and patents. To control for other factors affecting the probability of citation, the model also includes the citation lag (calculated using the publication year of both the cited article and citing patent), a set of country by cited year fixed effects (denoted $\mathbf{YC}_{i,t}$ in the equation below), which control for the different

[4] Low citation rates occur partially because many articles appear at the end of the sample, so there is less time for them to be cited. However, even for articles published earlier in the sample period, citation rates are less than ten percent (Popp 2016).

opportunities for future citations available to articles from different countries and from different times, and a dummy variable to control for articles with authors from multiple countries. As I explicitly model the effect of time using the citation lag, I use an exponential baseline hazard to model the probability of article *i* receiving a citation in year *t*:

(1) $h(i,t) = \exp(\alpha_0 + \alpha_1 citationlag + \alpha_2 citationlag^2 + \alpha_3 totalcitations + \alpha_4 multicountry + \boldsymbol{\gamma}\mathbf{YC_{i,t}})$

In each regression, standard errors are clustered by article.

While the intent of the country-by-year dummies is to control for changing citation opportunities, a possible concern is that the total citations received is nonetheless endogenous. For example, both alternative energy articles and patents are likely to appear shortly after major policy changes that promote alternative energy. As a result, opportunities for receiving citations both from other articles and from patents will increase. To address this possibility, I also consider a two-step procedure. First, I use a negative binomial regression to predict the total citations received by each article. Second, I replace *totalcitations* in equation (1) with a quality index derived from this regression, as described below.

To predict the total citations received by each article, I use a negative binomial regression that includes publication year fixed effects, a dummy variable for articles with authors from multiple countries, and the cumulative number of publications in the country of article *i*'s origin between the article's publication year and 2011. Since citations are more likely between articles from the same country, this value proxies for the citation opportunities available for article i.[5] For each article *i* from country *j* published in year *t*, I model total citations as:.[6]

(2) $totcites_{i,j,t} = f(\#domesticpublications_{j,t}, multicountry, \textbf{YEAR FIXED EFFECTS})$

[5] Note that the year fixed effects account for remaining citing opportunities from foreign publications.
[6] Note that I do not also include a cumulative count for publications from other countries, as these are controlled for by the year fixed effects.

where $\#domesticpublications_{j,t}$ represents the cumulative number of publications in the country of origin between the publication year t and 2011. There is a separate observation for each author country represented on an article. I use weighted regression to weight each observation by the share of authors from that country, so that articles with multiple authors do not receive more weight than other articles.

Using equation (2), for each article I then calculate the probability that the predicted citations are less than or equal to the actual citations received, which I multiply by 100 to create a *quality index* for each article.[7] This quality index can be interpreted similarly to an error term from a linear regression, with the advantage that it is bounded by 0 and 100. In the second step of my two-step procedure, I replace the actual count of total citations in equation (1) with this quality index.

Table 3 provides descriptive statistics on both the actual journal article citations receive and the quality index, which helps illustrate the intuition of the quality index. An article with more citations than expected given its characteristics (publication year and future citation opportunities from articles in the same country) will have a quality index close to 100, whereas an article receiving fewer citations than expected will receive a quality index close to 0. For each technology, the mean quality index is near 50. Actual citations are skewed, with a median value about half of the mean for each technology. Similarly, the median quality index is also lower than the mean, but by a much smaller amount. The table also shows actual citations and the quality index for various percentiles. At each percentile, the quality index is generally similar in value to that percentile (e.g. the median quality index is close to 50). Consistent with the skewed nature of the data, the 25th percentile of the quality index is slightly larger than 25, and the 75th percentile of the quality index is slightly lower than 75.

[7] Done using the command predict VARNAME, pr(0,total_ctes) in Stata.

[TABLE 3 ABOUT HERE]

Finally, to allow for possible non-linear effects of article quality on NPL citations, I also run regressions replacing the actual number of citations received (or the quality index) with a set of dummy variables for whether the actual citations (or quality index) are in the (1) 50^{th}-75^{th} percentile, (2) 75^{th}-90^{th} percentile, (3) 90^{th}-95^{th} percentile, or (4) 95^{th}-100^{th} percentile.

Table 4 presents results for the regressions using actual values, and Table 5 the results from regressions using percentile dummy variables. For the exponential hazard model, exp(β)-1 gives the marginal effect of either one additional citation (columns 1-3) or a one-unit increase in the quality index (columns 4-6). Table 4 shows that articles receiving more citations from other journal publications are also more likely to be cited by subsequent patents. An additional journal citation raises the probability of receiving a patent citation by 0.2% (solar) to 1.7% (wind). The larger magnitude for wind is primarily due to wind articles receiving fewer article citations on average than biofuels or solar. Similarly, a 10 percent increase in the quality index of an article (a bit less than one-half of a standard deviation of each index) increases the probability of receiving a patent citation by 22% (wind) to 33% (solar).

[TABLE 4 ABOUT HERE]

[TABLE 5 ABOUT HERE]

Table 5 shows that the effects of article quality on receiving a patent citation are non-linear. Compared to articles with citations or a quality index below the median, those in the 50^{th}-75^{th} percentile are generally 60-80% more likely to receive a patent citation, but results are statistically significant at a 5 percent level only for solar energy articles. Beyond the 75^{th} percentile, results are statistically significant at the five percent level for all technologies, with the exception of the quality index for wind articles. There, it is not until the 90^{th} percentile quality index that I observe a statistically significant impact on the probability of receiving a patent citation. Most importantly, in the highest percentiles, the increased probability of citation is large. Articles in the 90^{th}-95^{th} percentile for either actual publications or the quality index are 292-569 percent more likely to be cited by a patent. In the 95^{th}-100^{th} percentile, articles are 480-799% more likely to be cited by a patent.

These results suggest that the same articles that are perceived as important by other authors of journal publications (as indicated by larger citation counts from other articles) are also perceived important by inventors, who cite these articles more frequently on related patents. Given that many articles are never cited by patents, the skewed nature of the results in Table 5 are of particular interest, as it is indeed the highest quality articles (as perceived by other authors) that are also cited by patents. Since the ultimate goal of research on alternative energy sources is to develop new technologies, these results indicate that counts of journal-to-journal citations are a good indicator of the ultimate quality of an article.[8]

4. Knowledge flows across institutions

Having validated the use of citation data as a measure of article quality, I now use these citation data to examine both the quality of articles and patents across different research organizations and the flow of knowledge across these institutions. The key assumption is that institutions producing more widely cited research output are generating research of greater value

[8] A possible concern is that being cited by a patent changes the probability of being cited by other articles (Murray and Stern, 2007). To check for this possibility, I also ran regressions using the cumulative citations received through year *t*-1, rather than total citations received. The results for citations and the quality index were similar to those presented above, suggesting that changing citation behavior is not a problem in my sample. However, unlike the results presented here, the citation lag coefficients change when using predicted, rather than actual citations. Since cumulative citations are growing over time, this result suggests that using cumulative citations also picks up a time effect.

to future researchers.[9] I consider five types of research organizations: universities, governments (e.g. government laboratories), research institutes, private companies, and other organizations (including individual inventors).[10] To assess whether collaborations between different types of institutions encourage additional technology transfer, I classify collaborations between two or more institution types as a separate category. To further focus on technology transfer from the public to private sector, I distinguish between collaborations that include a private firm as one of the institutions from collaborations between two or more non-private institutions (e.g., academic and government). As organizations are identified based on author affiliations and patent assignees, the focus is on organizations *performing* research, not funding research. For example, government research only includes publications with government organization affiliates, such as U.S. Department of Energy laboratories. Research funded by governments and performed at universities are considered university research. The data provide evidence as to where public R&D funds can be targeted to best encourage the transfer of knowledge from more basic science to applied technology work.

4.1. Descriptive data on research institutions

Before evaluating citations between research institutions, it is important to know where both patented and published energy research comes from. Figures 3 and 4 show the breakdown of scientific articles and patents by organization. Appendix Figures A7-A12 provide separate breakdowns for US and foreign publications. Not surprisingly, universities account for most

---

[9] Jaffe, Fogarty, and Banks (1998) investigate the validity of this assumption, using evidence from patent-to-patent citations made to NASA patents. They conclude that, although there is noise in the citation process, aggregate citation patterns represent knowledge spillovers, although the spillover may be indirect. Jaffe and de Rassenfosse (2016) provide a recent review of the literature using patent citations as a proxy for the economics value of patented research.

[10] Organizations for articles are based on the affiliations of each author listed on the paper. Articles with multiple authors are weighted in the same way as described earlier for country affiliation. For patents, the organization is the patent assignee. Organization coding used a combination of text-based logic (e.g. “University of…”) and Internet searches to identify organization types not clear from the organization name.

scientific articles, providing between 60 and 66% of articles for each technology. Universities patent renewable energy technologies less frequently. However, university patents are relatively more prevalent for biofuels, perhaps indicating the importance of genetic research to new patentable biofuel technologies. 12.4% of biofuel patents are assigned to universities, compared to just 2.5% for solar and 1.5% for wind. As shown in Appendix Figure A8, this is almost entirely due to U.S. university patenting activity. 17.3% of biofuel patents assigned to US inventors come from universities, whereas just 3.6% of patents assigned to foreign inventors come from universities. Similarly, for both biofuels and wind (Table A11), while the share of US patents assigned to universities is somewhat lower than the share of foreign patents, the share of collaborative patents is higher. Technology transfer between research institutions appears more successful in the U.S. than elsewhere. This is consistent with earlier research on energy patents in the U.S. suggesting technology transfer improved after policy efforts such as the Bayh-Dole Act (encouraging university patenting) and the Federal Technology Transfer Act of 1980 (Popp 2006, Jaffe and Lerner 2001).

[FIGURE 3 ABOUT HERE]

[FIGURE 4 ABOUT HERE]

Private companies are the primary source for patents, although with a much larger range (55% for wind, compared to 72% for solar). In contrast, companies author just two to three percent of all scientific articles. However, collaboration is more common for scientific articles, as another six to nine percent of articles are collaborative efforts that include a private company author. Collaborations never make up more than 1.5% of patents for any technology.

The importance of government authors and inventors varies by technology. For scientific articles, government authors provide 3.8% (wind) to 6.4% (biofuels) of all articles. For patents, the range is much larger, with just 1.5% of wind patents assigned to the government, compared to 5.1% for solar and 11.6% for biofuels. Finally, for wind, just over 40% of patents are not assigned to any organization type. Most of these are patents assigned to an individual inventor.

Table 6 breaks down the share of articles receiving a patent citation by organization type. The left columns consider all citations, whereas the right columns exclude self-citations. Articles written by corporate authors are most likely to be cited by both solar and wind patents, even after removing self-citations. Government articles are next most likely to be cited by a patent, and are most likely to be cited by biofuel patents. Interestingly, despite being the most common source of articles, the share of university articles cited by a patent is always below the average for all articles, even after removing self-citations.

[TABLE 6 ABOUT HERE]

Finally, Tables 7 and 8 provide descriptive data on the number of article-article and patent-patent citations by organization type. For biofuels and wind, government articles receive the most citations per article. For solar, it is collaborations including a private sector author that receive the most. Interestingly, company articles receive the second most citations per article in solar energy, whereas company articles receive less citations than average in biofuels and wind. Not surprisingly, given the large number of university articles, citations per article for universities are near the overall mean for each technology. I observe similar trends for patents, with government patents receiving the most citations per patent for biofuels and wind, while company patents receive the most citations in solar energy.

[TABLE 7 ABOUT HERE]

[TABLE 8 ABOUT HERE]

4.2. Regression Analysis

While these descriptive data provide an initial indication of both the level of research activity and the importance of research emerging from different institutions, they do not control for other factors that may influence citations. For example, Appendix Figures A13-A18 show how the distribution of organizations has changed over time. For all three technologies, the share of university articles increases in the second half of the sample. As this is also the time of greatest publication activity, many university articles have thus had fewer opportunities to be cited than other articles. Thus, to study flows of knowledge across institutions, it is important to control for the opportunities a publication has to be cited.[11]

Following Jaffe and his co-authors (Caballero and Jaffe 1993, Jaffe and Trajtenberg 1996, 1999), I create groups of publications based on the year of publication, the organization(s) represented on the publication, and the country of origin. This allows me to study flows of knowledge across *pairs* of cited/citing publication categories. Using the subscripts *CTD* and *CTG* to represent the cited and citing cohorts, respectively, the probability of citation, *p*, for publications within each citing/cited cohort pair is:

$$p_{CTD,CTG} = \frac{c_{CTD,CTG}}{(n_{CTD})(n_{CTG})}. \tag{3}$$

Here, n represents the number of publications in a cohort and c the total number of citations between publications in each cohort pair.

To estimate the likelihood of citation for various groups of publications, I use the model developed by Jaffe and his co-authors to control for factors affecting the likelihood of citation. The probability that a publication in potentially cited cohort *CTD* receives a citation from a

[11] For ease of exposition, I use “publication” to refer to either the publication of a US patent or of a scientific article, as the citation model is used to analyze both article and patent citations.

publication in potentially citing cohort *CTG* is a function of the characteristics of the citing and cited publication groups, $\alpha(CTD,CTG)$, and the time that has passed since the initial publication. Using β_1 to represent the rate of decay of knowledge as it becomes obsolete, and β_2 for the rate at which newly produced knowledge diffuses, the probability of citation is written as:

(4) $p(CTG,CTD) = \alpha(CTD,CTG)\exp[-\beta_1(t_{CTG}-t_{CTD})][1-\exp(-\beta_2(t_{CTG}-t_{CTD}+1))]+\varepsilon.$

$\alpha(CTD,CTG)$ represents the effect of various publication cohort characteristics on the citation probability. Adding one to the lag between citing and cited patents for the rate of diffusion ensures that patents can possibly be cited in the first year. I control for the following:

- the technology category (biofuels, solar, or wind)
- the organization of both the cited and citing publication
- the country of origin of the cited and citing publication[12]
- the year of publication of the citing cohort[13]

For example, one possible citing-cited cohort includes citations made by US solar energy patents assigned to corporations and granted in 2010 to US solar energy patents assigned to universities and granted in 2005. Another such cohort would include citations made by US solar energy patents assigned to the government and granted in 2008 to Japanese solar energy patents assigned to corporations and granted in 2000.

In this paper, the effect of the citing and cited organization is of primary interest. Note that α enters the model multiplicatively, so that the null hypothesis of no effect corresponds with

[12] The countries considered vary by publication type to focus on those countries most prevalent in the data. For patents, country categories include Canada, Japan, the U.S., Denmark, Germany, Great Britain, Spain, other European Patent Office members, and all other countries. For articles, country categories are Canada, China, Japan, the US, Denmark, Germany, Spain, Turkey, other European Patent Office members, and all other countries. For patents citing articles, I only distinguish between US and foreign articles and patents. Due to the smaller number of article/patent citations, most cohort pairs include zero citations if finer distinctions of the data are used.

[13] As the model also includes the lag between citing and cited documents, I do not separately include the publication year of the cited document, since that is a function of the lag and the citing year.

a value of 1, not 0. For purposes of identification, one α parameter of each type is normalized to 1. Thus, for any cohort characteristic, an estimate of $\alpha(CTD,CTG)$ greater than one means a citation is more likely (compared to the normalized base category) when a publication with characteristic j is part of the citing/cited cohort. Similarly, an estimate of $\alpha(CTD,CTG)$ less than one means a citation is less likely. I estimate equation (4) using non-linear least squares. Because the data are grouped, I weight each observation by $\sqrt{(n_{CTD})(N_{CTG})}$ to avoid problems with heteroskedasticity.

[TABLE 9 ABOUT HERE]

I estimate separate models for article-to-article, patent-to-patent, and patent-to-article citations. Table 9 shows the result of my main specification. Columns 1-3 provide results excluding self-citations, and columns 4-6 include self-citations. Column 1 shows results for article-to-article citations, column 2 for patent-to-patent citations, and column 3 for patent-to-article citations. The pattern repeats for columns 4-6. In all cases, university articles or patents are the base category. Recall that, since the model is multiplicative, the base category is normalized to one. For example, looking at cited institutions, the coefficient of 0.858 for government articles in column 1 means that government articles are 14.2 percent less likely to be cited than university articles. Similarly, in column 2, government patents are 13.6 percent more likely to be cited than university patents. In general, the results with and without self-citations are similar. The main exception is for patents citing scientific articles (columns 3 and 6). Except for collaborations, the coefficients for cited organizations fall between 25 and 50 percent when including self-citations. Since this is relative to the base category of university articles, this suggests that university patent-to-publication citations are more likely to be self-citations. Since my main interest is knowledge flows, I focus on the results excluding self-citations below except when exceptions arise.

Looking first at articles (columns 1 & 4), university articles are both more likely to be cited and are more likely to cite other research, as the coefficients for all other organizations are less than one. This is consistent with the notion that most basic research comes from a university setting.[14] The one exception is that company collaborations are 3 percent more likely to be cited, but only when self-citations are included. Similarly, when including self-citations, non-company collaborations are 2.1 percent more likely to cite other work. As these results only hold for self-citations, they suggest collaborations help expose the research partners to each other's articles.

The relative importance of institutions changes when looking at patent-to-patent citations (columns 2 & 5). I find two important results here. First, research performed by government institutions is highly valuable. Government patents are 13.6 percent more likely to be cited than university patents. Second, collaborative research enhances the flow of knowledge across institutions. Non-company collaborations are 30 percent more likely to be cited than university patents, although the result is only significant at the ten percent level. Note that this excludes self-citations between any of the participating organizations. Thus, rather than it simply being the case that collaborations enhance technology transfer within the group, this result suggests that the patents resulting from collaborations make novel contributions that are more valuable to future researchers. Indeed, note that these collaborations are not more likely to be citing other patents, further emphasizing that it is not simply an increased propensity for cross-citation

[14] Note that this differs from the raw descriptive data, where university articles received fewer citations on average than government articles. As government articles were more prevalent early in the sample, they have had more opportunities for citation, thus illustrating the need to control for citation opportunities to properly assess citation counts across organization types.

generating this result. Moreover, note that patents from other institutions are no less likely to be cited than university patents.

The primacy of research performed in government laboratories continues when looking at patent-to-article citations (columns 3 & 6). Government articles are 14 percent more likely to be cited by a patent than other research. Thus, government research does appear to play an important translational role linking basic and applied research. However, university articles are still important, as they are cited more frequently than non-government articles. Finally, when excluding self-citations, non-company collaborations are 60 percent more likely than university patents to cite other articles, providing further evidence that these collaborations expose research partners to a wider range of knowledge than they would obtain on their own.

Separating the results by technology yields a few interesting differences. As the results are similar with and without self-citations, I present only the results excluding self-citations in Table 10.[15] For article-to-article citations in biofuels and solar energy, university articles generally remain the most valuable. However, company collaboration articles are 10 percent more likely to be cited than university articles. For wind energy, things are quite different. Here, company and government articles are more important. Company articles are 62 percent more likely to be cited, and government articles are 80 percent more likely to be cited. Similarly, both company and non-company collaborations are 115 percent and 74 percent more likely to be cited, respectively. Wind is a more mature technology and is approaching cost competitiveness with fossil fuels in ideal locations. As such, these results suggest that wind research is moving towards a more applied stage where university research becomes less significant.

[TABLE 10 ABOUT HERE]

[15] Because of the smaller number of patent-to-article citations, that model does not converge for individual technologies.

As before, for patent-to-patent citations, government biofuel and wind patents are more likely be cited. However, government solar energy patents are 10 percent less likely to be cited than university patents. Emphasizing the role of the private sector, company patents in biofuels and solar energy are 9 and 22 percent more likely to be cited than university patents. Company patents in these two technologies are also 70-75 percent more likely to cite other patents.[16]

Finally, Figure 5 illustrates the impact of decay and diffusion. It shows how the probability of citation varies over time for articles, patents, and patent-to-article citations. The probability of article-to-article citations peaks first, just 3 years after publication.[17] In contrast, patent-to-patent citations are more durable. While the probabilities for article and patent citations start at similar levels, the probability of patent citations continues to increase for up to 9 years after the patent was granted.[18] Note that the probability of citation does not fall as quickly for patents as it does for articles. Finally, knowledge flows between articles and patents are both less frequent and take longer to occur. For patents citing articles, the probability peaks 15 years after the article was published.[19]

[FIGURE 5 ABOUT HERE]

---

[16] The regressions also include controls for the citing and cited countries. While they are important controls, they provide less interesting information. Both US articles and patents are the most commonly cited. The next most commonly cited countries are other English-speaking nations: the UK and Canada. Articles and patents from most other countries are cited about 30-40 percent less frequently than US articles. Within scientific articles, Chinese publications are 13 percent more likely to have citations, and Turkish articles are 27 percent more likely to have citations, suggesting work from those countries may be more derivative of earlier research. Finally, since the patent data includes only patents granted in the US, not surprisingly US scientific articles are cited about three times as often on these patents as are foreign articles. Appendix Table A3 provides complete regression results for the pooled model excluding self-citations, including country coefficients.

[17] One comparable study looking at citations across publications in other fields is Adams and Clemmons (2013). They find slightly longer mean citation lags of about 6 years. However, they find the fastest lags within physics, a field very relevant for energy research. There, the mean lag is a similar 3.5 years.

[18] Popp *et al*. (2013) finds similar duration for energy patents in a study looking at differences across energy technologies. Patent citations to energy patents appear more durable than for other technologies, as most previous studies find citation probabilities peaking 3-4 years after grant (e.g. Jaffe and Trajtenberg, 2002)

[19] For comparison, in a broader study of NPL citations, Branstatter and Ogura (2005) find that patent citations to scientific publications begin shortly after article publication, and that the probability of citation peaks about eight years after article publication.

4.3. Discussion

The regression results have several possible implications for energy R&D policy. While a full discussion of the ideal role for government energy R&D is beyond the scope of this paper, here I outline some general guidelines based on the existing literature and discuss how my results inform future energy R&D decisions. First, earlier studies do find instances where government energy R&D crowds out private R&D efforts, particularly when government funding targets applied research topics (Popp 2002). If one of the objectives of government policy is to avoid duplicating and potentially crowding out private research efforts, government R&D will be most effective if it focuses on breakthrough technologies that are not yet close to market.

With this in mind, the results for wind energy suggest that additional public R&D support for wind is of less value than for solar or biofuels. For wind energy, the most valuable scientific articles are coming further downstream in the research process. Government and company articles are more frequently cited than university research, as are multi-institution collaborations. While my data do not allow me to identify public R&D support that these private sector research projects may have received, the results demonstrate that wind energy R&D has reached a more mature phase, where the focus should be on product development, rather than basic scientific advances. Much of this work will be carried out in the private sector. Recall from Figure 1 that the greatest growth in patenting among the technologies in this paper was in wind, and that wind patents increased at a faster rate than wind articles. To create demand for wind energy innovation, other policy mechanisms will need to be in place, such as carbon pricing or renewable energy mandates (e.g. Popp, 2010). However, if governments wish to avoid duplicating private R&D efforts, public energy R&D investments would be better spent on technologies that complement private sector wind energy investment, such as enhancing energy storage, than on duplicating what the private sector will do on its own.

Government research can also help new technologies overcome roadblocks to commercialization (Mowrey *et al.* 2010, Weyant 2011). Nonetheless, government R&D is often portrayed as wasteful and inefficient in public policy debates. Cohen and Noll's (1991) *Technology Pork Barrel* provides the classic academic treatise on this viewpoint, arguing that political realities influence funding decisions in ways that may prolong unsuccessful projects. However, I show here that research on renewable energy sources produced by government institutions has been particularly helpful moving alternative energy research to an applied stage. Patents assigned to government research institutions more likely to be cited than those from other institutions. More importantly, scientific articles from government institutions are more likely than other articles to be cited by patents. Although government articles are not more likely to be cited by other scientific articles, these government articles appear to be important links between basic and applied research. Further research to uncover the value of the patents inspired by government research would be a useful next step.

Finally, the long lags between patent-article citations pose a political challenge. Policymakers face political constraints making it difficult to support policies with little short term payoff. While ideally government funded R&D funding should focus on riskier projects less likely to be performed in the private sector, the long lags between an initial publication and the ultimate technology development from such projects may make it difficult to sustain political support for research on these long-term projects. Thus, a second-best solution may be for governments to develop a diverse portfolio of projects that includes some low-risk projects likely to have relatively quick returns. While these may result in some crowding out of private R&D,

such success stories will help build public support for a continuous, steady stream of public energy R&D funding. Funding agencies will need to weigh the cost of such crowding out against the potential gains of political support for a portfolio of research that also includes the necessary risker, but less politically popular, R&D projects.

5. Conclusions

Expansion of government energy R&D budgets is likely to continue to be a key component of climate policy. Using an original data set of both scientific articles and patents pertaining to alternative energy technologies, this paper provides new evidence on the flows of knowledge between university, private, and government research. The paper makes three key contributions to both the study of energy research and to energy R&D policy.

First, using scientific articles to represent more basic, upstream research and patents to represent more applied downstream research, the paper provides new descriptive data on the flows of basic and applied research across institutions. There are notable differences in the importance of different institutions across technologies. For example, university patents are relatively more prevalent for biofuels, perhaps indicating the importance of genetic research to new patentable biofuel technologies. Similarly, wind research has moved to a more applied stage, where the most valuable inventions come from downstream institutions such as the private sector, rather than from universities.

Second, using non-patent literature references to link articles and patents, I show that highly cited academic literature is also valuable to the creation of applied technology (e.g. patents). The results are highly skewed. Few articles are ever cited by patents, and those that are cited by patents are amongst the most highly cited articles by other research articles. This result indicates that not only are journal article citations a good indicator of the usefulness of the research to other academic researchers, but also provide information on the value of published research for potential technology development.

Finally, analysis of citation flows across institutions highlights the high value of research performed at government institutions. Patents assigned to governments are more likely to be cited than other patents. Moreover, government articles are more likely to be cited by patents than any other institution, including universities. Thus, research performed at government institutions appears to play an important translational role linking basic and applied research. Funding agencies may wish to expand the role of government research facilities as they increase public energy R&D budgets.

References

Adams, J. D. &Clemmons 2013. How rapidly does Science leak out? A Study of the Diffusion of Fundamental Ideas. *Jounral of Human Capital*, 7(3): 191-223.

Branstatter, Lee and Yoshiaki Ogura. 2005. Is academic science driving a surge in industrial innovation? Evidence from patent counts. *NBER Working Paper #11561*.

Caballero, R.J. and A. B. Jaffe. 1993. How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In: O.J. Blanchard and S. Fischer (Eds.), *NBER Macroeconomics Annual 1993*, MIT Press, Cambridge, MA, 15-74.

Cohen, L.R. and R.G. Noll. 1991. *The Technology Pork Barrel*, Brookings Institute, Washington, DC.

Davenport, Coral and Nick Wingfield. 2015. The Billionaire Stateman: Gates is showing how business leaders can influence world leaders to take action on tackling climate change. *The New York Times*, A1 & A7.

Dechezleprêtre, A., Glachant, M. 2014. Does foreign environmental policy influence domestic innovation? Evidence from the wind industry. Environmental and Resource Economics 58: 391-413.

Dechezleprêtre, A., Popp, D. forthcoming. Fiscal and regulatory instruments for clean technology deployment in the European Union. In Ian Parry, Karen Pittel, and Herman Vollebergh (eds.) *Energy Tax and Regulatory Policy in Europe: Reform Priorities*. MIT Press, Cambridge, MA.

Greenstone, Michael, and Adam Looney. 2012. Paying Too Much for Energy? The True Costs of Our Energy Choices. *Dædalus* 141(2): 10–30.

IEA 2014. International Energy Agency, *CO_2 Emissions from Fuel Combustion: Highlight*s, OECD/IEA, Paris, France.

Jaffe, A.B. and de Rassenfosse, G. 2016. Patent citation data in social science research: Overview and best practices. *NBER Working Paper #21868*.

Jaffe, A.B., Fogarty, M.S. and Banks, B.A. 1998. Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation. *Journal of Industrial Economics*, 46(2): 183–205.

Jaffe, A.B. and J. Lerner. 2001. Reinventing public R&D: patent policy and the commercialization of national laboratory technologies. *RAND Journal of Economics*, 32(1), 167-198.

Jaffe, A.B., and M. Trajtenberg, International knowledge flows: Evidence from patent citations, *Economics of Innovation and New Technology*. 8 (1999) 105-136.

Jaffe, A.B., and M. Trajtenberg, Flows of knowledge from universities and federal labs: Modeling the flow of patent citations over time and across institutional and geographic boundaries, *Proceedings of the National Academies of Sciences* 93 (1996) 12671-12677.

Jaffe, A.B. and Trajtenberg, M. 2002. *Patents, Citations, & Innovations: A Window on the Knowledge Economy,* MIT Press, Cambridge, MA.

Johnstone, Nick, Ivan Haščič, and David Popp. 2010. Renewable Energy Policies and Technological Innovation: Evidence Based on Patent Counts. *Environmental and Resource Economics* 45(1): 133–55.

Mowrey, D.C., R.R. Nelson, and B.R. Martin. 2011. Technology policy and global warming: Why new policy models are needed (or why putting new wine in old bottles won't work). *Research Policy* 39: 1011-1023.

Murray, Fiona and Scott Stern.2007. When Ideas Are Not Free: The Impact of Patents on Scientific Research. in *Innovation Policy and the Economy vol. 7*, Adam B. Jaffe, Josh Lerner, and Scott Stern (eds.), MIT Press, Cambridge, MA, 33-69.

Nesta, L., F. Vona, and F. Nicolli. 2014. Environmental policies, competition, and innovation in renewable energy. *Journal of Environmental Economics and Management*, 67, 396-411.

Peters, M., Schneider, M., Griesshaber, T., Hoffman, V.H. (2012), The impact of technology-push and demand-pull policies on technical change – Does the locus of policies matter? *Research Policy* 41: 1296-1308.

Popp, David. 2016. Economic Analysis of Scientific Publications and Implications for Energy Research and Development. *Nature Energy* 1(4): 1–8.

Popp. D. 2010. Innovation and Climate Policy. In *Annual Review of Resource Economics*, vol. 2., Gordon C. Rausser, V. Kerry Smith and David Zilberman eds., Annual Reviews, Palo Alto, CA, 275-298.

Popp, D. 2006. They Don't Invent Them Like They Used To: An Examination of Energy Patent Citations Over Time. *Economics of Innovation and New Technology* 15 (8): 753-776.

Popp, D. 2002. Induced innovation and energy prices. *American Economic Review* 92(1), 160-180.

Popp , D., N. Santen, K. Fisher-Vanden, and M. Webster. 2013. Technology variation vs. R&D uncertainty: What matters most for energy patent success? *Resource and Energy Economics* 35(4): 505-533.

Roach, M. & Cohen. 2013. W.M. Lens or Prism? Patent Citations as a Measure of Knowledge Flows from Public Research. *Management Science* 59(2), 504-525.

Sorenson, Olav and Lee Fleming. 2004. Science and the diffusion of knowledge. *Research Policy*, 33: 1615-1634.

Verdolini, Elena, and Marzio Galeotti. 2011. At Home and Abroad: An Empirical Analysis of Innovation and Diffusion in Energy Technologies. *Journal of Environmental Economics and Management* 61: 119–34.

Veugelers, R. 2012. Which policy instruments to induce clean innovating? *Research Policy* 41: 1770-1778.

Weyant, J. 2011. Accelerating the development and diffusion of new energy technologies: Beyond the "valley of death." *Energy Economics* 33: 674-682.

Figure 1: Total patents

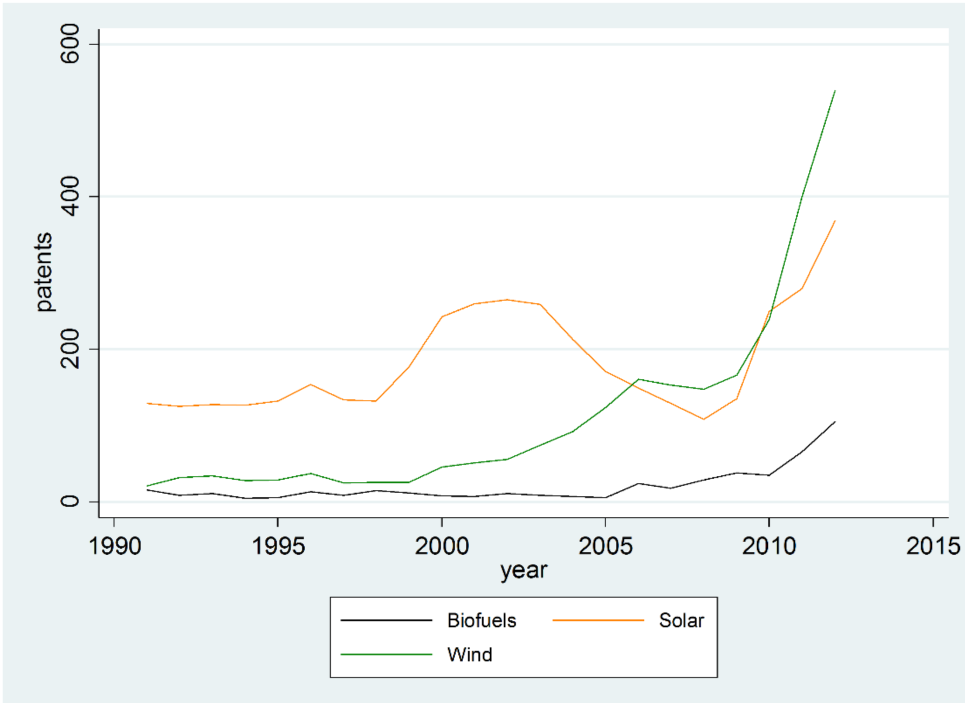

Figure shows total patents for clean energy technologies granted by the USPTO in each year.

Figure 2: Total scientific articles

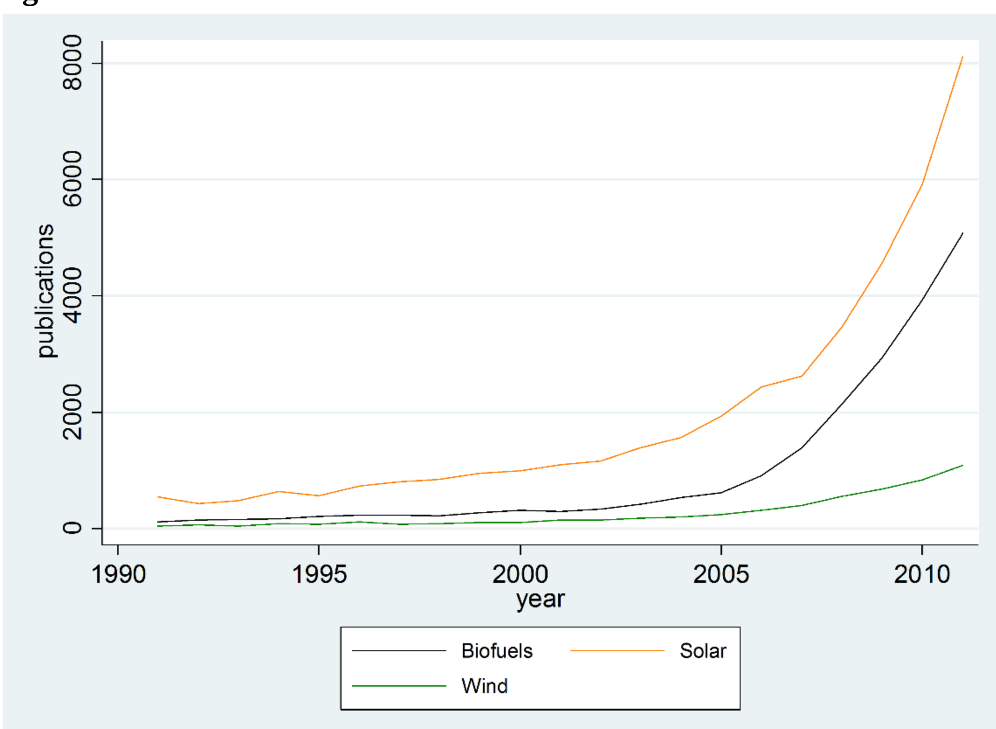

Figure shows the number of scientific articles for clean energy technologies published by year.

Figure 3: Article author organizations

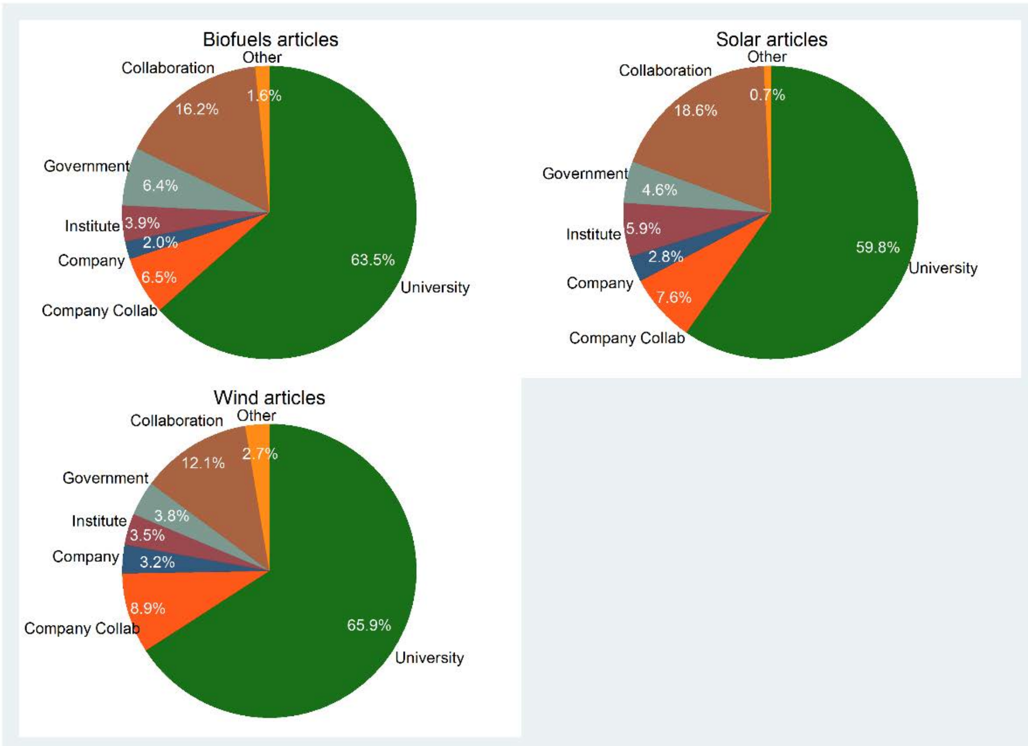

Figure 4: Patent inventor organizations

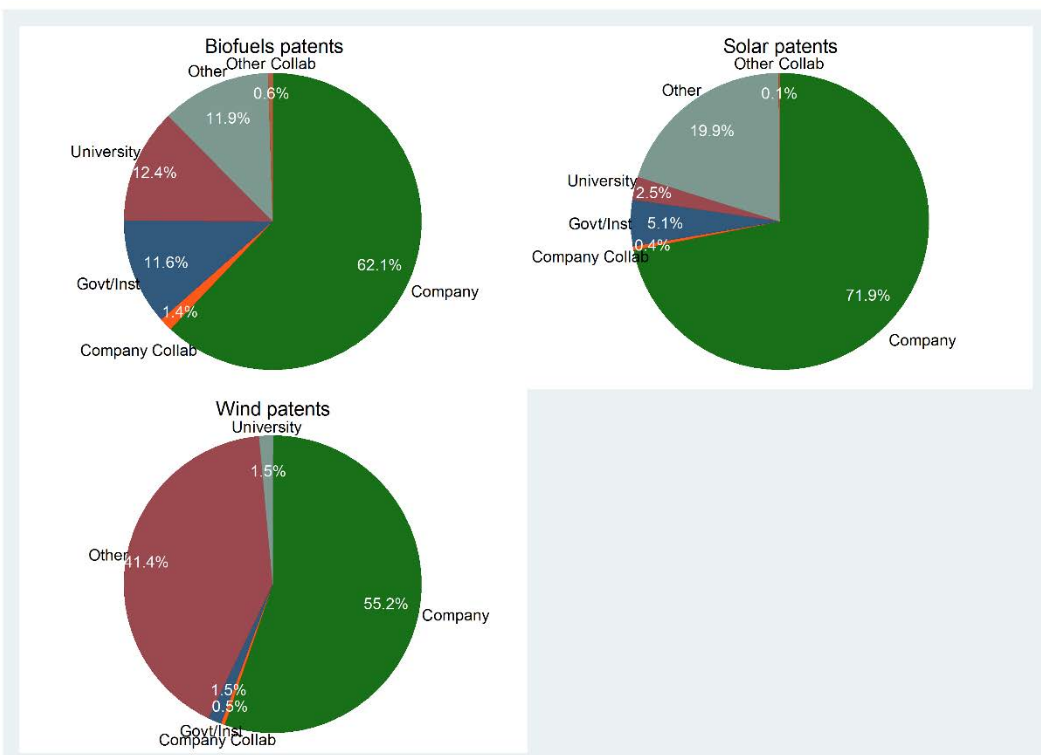

Figure 5 – Probability of citation over time

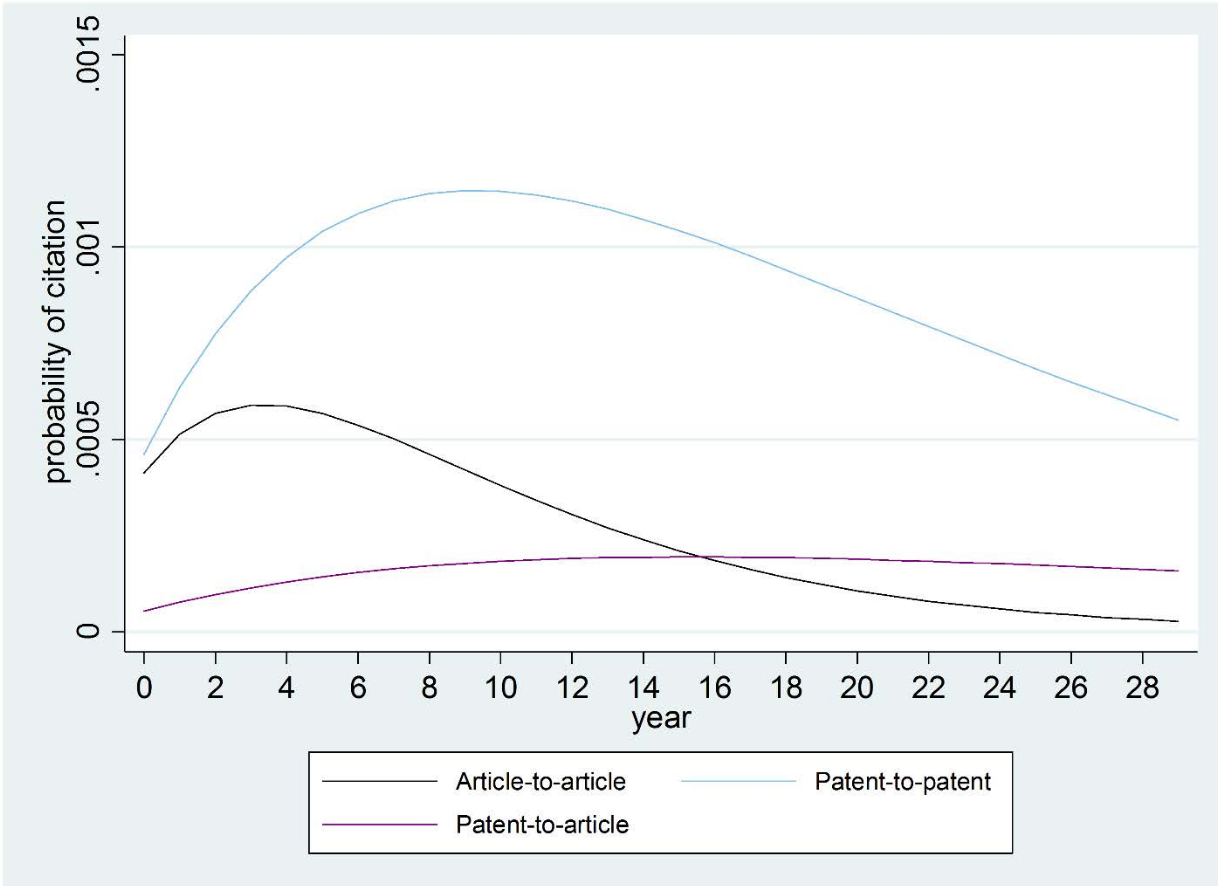

Figure shows the baseline probability of citation for different publication combinations over time, based on the estimated rates of decay and diffusion excluding self-citations in Table 9. The *x*-axis plots the years between the publication dates of the citing and cited documents.

Table 1 – Top energy article and patent countries

Top 10 article sources, weighted publication counts

Biofuels		**Solar Energy**		**Wind**	
United States	4428.62	United States	6323.95	United States	916.47
Peoples R China	1817.09	Peoples R China	5044.19	United Kingdom	571.13
India	1279.08	Japan	4314.37	Denmark	337.15
Brazil	949.75	Germany	3525.28	Germany	290.77
Turkey	793.54	South Korea	2415.65	Spain	268.43
Japan	761.61	India	2123.81	Peoples R China	255.60
United Kingdom	749.06	Taiwan	1506.67	Canada	251.87
Canada	735.17	United Kingdom	1409.72	Japan	222.81
Germany	735.05	France	1219.91	Greece	200.63
Spain	714.40	Spain	1187.10	Turkey	197.62

Top 10 patent sources, weighted patent counts

Biofuels		**Solar Energy**		**Wind**	
United States	227.71	United States	1643.49	United States	892.05
Denmark	16.81	Japan	1196.33	Germany	340.43
Canada	15.17	Germany	310.10	Japan	135.50
Netherlands	13.81	Australia	71.40	Denmark	129.67
Japan	13.57	Taiwan	53.67	Canada	79.52
Germany	9.83	France	49.57	Taiwan	53.33
Great Britain	9.00	Canada	46.58	Great Britain	47.28
Finland	8.46	Switzerland	44.19	Spain	31.92
India	7.25	Israel	40.47	Netherlands	31.88
France	5.00	South Korea	39.00	France	27.50

Table 2 –Citation summary statistics

	N	mean	sd	min	p25	p50	p75	p90	p95	max
Total article/article citations										
Biofuels	20649	10.16	26.41	0	0	3	10	26	42	1561
Solar	41276	12.86	58.04	0	1	4	11	28	48	7135
Wind	5615	5.56	11.06	0.000	0.000	2.000	7.000	15.000	23.000	273
Total patent/patent citations										
Biofuels	354	1.56	3.26	0	0	0	2	5	8	23
Solar	3700	2.71	4.94	0	0	1	3	8	13	51
Wind	1968	3.71	6.81	0	0	1	5	10	16	95
Total article/patent citations										
Biofuels	20649	0.01	0.15	0	0	0	0	0	0	6
Solar	41276	0.03	0.39	0	0	0	0	0	0	39
Wind	5615	0.02	0.31	0	0	0	0	0	0	11
Total article/patent citations: positive citations only										
Biofuels	136	1.51	0.98	1	1	1	2	3	3	6
Solar	540	2.13	2.67	1	1	1	2	4	7	39
Wind	64	1.97	2.12	1	1	1	2	3	5	11

Table 3 – Article citations and quality index

	N	mean	sd	min	p25	p50	p75	p90	p95	max
Total citations										
Biofuels	13721	17.399	34.223	0	3	9	20	40	60	1561
Solar	33465	18.924	68.507	0	3	7	18	40	68	7135
Wind	4403	8.312	12.996	0	1	4	11	21	30	273
Quality index										
Biofuels	13721	47.710	24.852	2.845	26.935	44.921	66.964	84.097	91.611	100
Solar	33465	48.641	23.210	7.675	29.619	45.755	64.954	83.542	91.901	100
Wind	4403	51.665	23.659	6.870	32.889	50.267	70.258	85.602	91.787	100

Table includes descriptive statistics for scientific article citations used in the regression for equation (2), as well as the resulting quality index.

Table 4 – Are highly cited journal articles also cited by patents?

	(1)	(2)	(3)	(4)	(5)	(6)
		Total Citations			*Predicted citations probability*	
	Biofuels	Solar	Wind	Biofuels	Solar	Wind
citation lag	0.655***	0.336***	0.470***	0.654***	0.329***	0.470***
	(0.135)	(0.0534)	(0.157)	(0.136)	(0.0523)	(0.156)
(citation lag)^2	-0.0241***	-0.0166***	-0.0145*	-0.0242***	-0.0164***	-0.0144*
	(0.00735)	(0.00365)	(0.00751)	(0.00743)	(0.00359)	(0.00740)
Total citations	0.00582***	0.00244***	0.0165***			
	(0.00172)	(0.000313)	(0.00580)			
Predicted citations probability				0.0237***	0.0320***	0.0216***
				(0.00501)	(0.00215)	(0.00615)
Multiple country dummy	0.111	-0.0485	-0.998	0.208	0.0220	-1.128*
	(0.344)	(0.133)	(0.630)	(0.332)	(0.134)	(0.664)
N	80316	239265	29498	80316	239265	29498
log likelihood	-294.7	-1915.3	-149.7	-291.8	-1840.1	-146.3

* p<0.10, ** p<0.05, *** p<0.01

Standard errors in parentheses. All standard errors clustered by article. Weighted regression using country weights for multicountry articles.

Table 5 – Are highly cited journal articles also cited by patents? Regressions with percentile dummies

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Citations			*Predicted citations probability*		
	Biofuels	Solar	Wind	Biofuels	Solar	Wind
citation lag	0.654***	0.330***	0.473***	0.655***	0.330***	0.469***
	(0.136)	(0.0523)	(0.156)	(0.136)	(0.0523)	(0.158)
(citation lag)^2	-0.0242***	-0.0164***	-0.0144*	-0.0242***	-0.0164***	-0.0142*
	(0.00745)	(0.00359)	(0.00741)	(0.00744)	(0.00359)	(0.00745)
50th-75th percentile	0.478	0.679***	0.903*	0.585*	0.573***	0.340
	(0.413)	(0.146)	(0.480)	(0.315)	(0.148)	(0.473)
75th-90th percentile	0.904**	1.088***	1.367***	0.780**	1.233***	0.599
	(0.374)	(0.152)	(0.446)	(0.350)	(0.146)	(0.465)
90th-95th percentile	1.536***	1.737***	1.900***	1.628***	1.769***	1.365**
	(0.416)	(0.174)	(0.593)	(0.423)	(0.169)	(0.680)
95th-100th percentile	1.826***	2.166***	1.906***	1.942***	2.196***	1.758***
	(0.387)	(0.153)	(0.519)	(0.354)	(0.150)	(0.455)
Multiple country dummy	0.100	-0.133	-1.326*	0.220	0.0244	-1.164*
	(0.335)	(0.133)	(0.685)	(0.336)	(0.134)	(0.664)
N	80316	239265	29498	80316	239265	29498
log likelihood	-290.0	-1855.0	-143.5	-289.5	-1845.7	-145.9

* p<0.10, ** p<0.05, *** p<0.01

Standard errors in parentheses. All standard errors clustered by article. Weighted regression using country weights for multicountry articles.

Table 6 – Percentage of articles cited by patents, by organization and country

A. Biofuels

	All citations			No self-citations		
	Foreign	USA	Total	Foreign	USA	Total
Collaboration	0.33%	1.29%	0.60%	0.33%	0.97%	0.51%
Company	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Company Collab	0.99%	2.65%	1.43%	0.99%	2.65%	1.43%
Government	0.25%	4.13%	1.82%	0.25%	3.94%	1.74%
Institute	0.26%	4.81%	0.49%	0.00%	4.81%	0.25%
Other	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
University	0.37%	1.23%	0.53%	0.36%	1.00%	0.48%
Total	0.38%	1.70%	0.66%	0.36%	1.48%	0.60%

B. Solar energy

	All citations			No self-citations		
	Foreign	USA	Total	Foreign	USA	Total
Collaboration	0.74%	2.06%	0.90%	0.68%	2.06%	0.85%
Company	3.34%	6.64%	4.18%	2.29%	5.02%	2.99%
Company Collab	1.45%	2.30%	1.63%	1.07%	2.08%	1.28%
Government	1.05%	3.70%	1.90%	1.05%	3.70%	1.90%
Institute	1.85%	0.00%	1.81%	1.60%	0.00%	1.56%
Other	0.81%	1.69%	0.98%	0.81%	1.69%	0.98%
University	1.09%	1.62%	1.17%	1.02%	1.48%	1.09%
Total	1.15%	2.18%	1.31%	1.03%	2.00%	1.18%

C. Wind energy

	All citations			No self-citations		
	Foreign	USA	Total	Foreign	USA	Total
Collaboration	0.76%	2.58%	1.17%	0.57%	2.58%	1.03%
Company	3.96%	3.93%	3.95%	3.96%	3.93%	3.95%
Company Collab	0.25%	3.17%	0.81%	0.25%	2.11%	0.60%
Government	1.38%	7.10%	3.26%	1.38%	7.10%	3.26%
Institute	0.55%	0.00%	0.51%	0.55%	0.00%	0.51%
Other	0.77%	0.00%	0.67%	0.77%	0.00%	0.67%
University	0.91%	1.36%	0.97%	0.91%	1.36%	0.97%
Total	0.92%	2.29%	1.14%	0.89%	2.18%	1.10%

Table 7 – Article citations by technology and organization type

Biofuels	N	mean	sd	min	p25	p50	p75	p90	p95	max
Collaboration	3349	10.54	26.16	0	0	3	11	26	43	625
Company	409	8.46	17.48	0	0	2	8	27	40	169
Company Collab	1332	10.49	22.52	0	0	3	11	26	43	349
Government	1320	13.28	25.16	0	1	5	15	34.5	52.5	288
Institute	811	8.39	16.14	0	0	2	9	22	40	146
Other	324	6.79	14.28	0	0	1	7	20	31	140
University	13104	9.97	27.87	0	0	3	10	25	41	1561
Total	20649	10.16	26.41	0	0	3	10	26	42	1561
Solar Energy	**N**	**mean**	**sd**	**min**	**p25**	**p50**	**p75**	**p90**	**p95**	**max**
Collaboration	7663	12.04	35.04	0	1	4	11	28	47	1727
Company	1172	14.11	30.25	0	2	5	15	32	55	484
Company Collab	3130	15.59	48.90	0	1	5	13	31	57	1347
Government	1898	13.88	35.52	0	1	5	13	29	52	510
Institute	2432	10.25	25.30	0	1	3	10	25	43	631
Other	307	5.15	10.48	0	0	2	6	13	21	64
University	24674	12.98	68.88	0	1	3	11	27	48	7135
Total	41276	12.86	58.04	0	1	4	11	28	48	7135
Wind Energy	**N**	**mean**	**sd**	**min**	**p25**	**p50**	**p75**	**p90**	**p95**	**max**
Collaboration	682	5.56	9.37	0	0	2	7	16	23	126
Company	177	4.80	11.36	0	0	1	5	15	22	124
Company Collab	497	4.82	9.19	0	0	1	6	13	22	89
Government	215	7.76	15.35	0	0	3	8	21	33	156
Institute	195	5.13	8.87	0	0	2	6	14	20	56
Other	150	4.18	7.19	0	0	1	5	13	19	40
University	3699	5.65	11.48	0	0	2	7	16	24	273
Total	5615	5.56	11.06	0	0	2	7	15	23	273

Table 8 – Patent citations by technology and organization type

Biofuels	**N**	**mean**	**sd**	**min**	**p25**	**p50**	**p75**	**p90**	**p95**	**max**
Company	220	1.30	2.70	0	0	0	2	4	7	19
Company Collab	5	0.00	0.00	0	0	0	0	0	0	0
Govt/Inst	41	3.34	5.38	0	0	1	4	11	15	23
Other	42	1.19	3.10	0	0	0	1	4	7	16
Other Collab	2	0.00	0.00	0	0	0	0	0	0	0
University	44	1.80	3.12	0	0	0	2	8	9	11
Total	354	1.56	3.26	0	0	0	2	5	8	23
Solar Energy	**N**	**mean**	**sd**	**min**	**p25**	**p50**	**p75**	**p90**	**p95**	**max**
Company	2662	2.86	5.19	0	0	1	3	9	13	51
Company Collab	14	0.21	0.58	0	0	0	0	1	2	2
Govt/Inst	188	1.79	2.92	0	0	1	3	5	9	17
Other	738	2.57	4.49	0	0	1	3	8	12	42
Other Collab	5	1.00	1.73	0	0	0	1	4	4	4
University	93	2.16	4.52	0	0	0	2	6	13	27
Total	3700	2.71	4.94	0	0	1	3	8	13	51
Wind Energy	**N**	**mean**	**sd**	**min**	**p25**	**p50**	**p75**	**p90**	**p95**	**max**
Company	1087	3.19	6.92	0	0	1	4	9	14	95
Company Collab	9	3.56	3.36	0	0	3	6	9	9	9
Govt/Inst	29	5.93	11.46	0	0	1	8	22	24	55
Other	814	4.31	6.34	0	0	2	6	12	17	44
University	29	4.07	8.46	0	0	1	4	10	10	45
Total	1968	3.71	6.81	0	0	1	5	10	16	95

Table 9 – Knowledge flows across organizations

	No self-citations			*Including self-citations*		
	art-art	**pat-pat**	**pat-art**	**art-art**	**pat-pat**	**pat-art**
Cited parameters						
Company	0.728***	1.041	0.532***	0.702***	1.094***	0.354***
	(0.009)	(0.027)	(0.081)	(0.010)	(0.028)	(0.071)
Company Collab.	1.005	0.832	0.648***	1.031***	0.865	0.489***
	(0.006)	(0.109)	(0.068)	(0.007)	(0.107)	(0.060)
Government	0.858***	1.136***	1.142**	0.896***	1.141***	0.699***
	(0.007)	(0.037)	(0.063)	(0.007)	(0.038)	(0.049)
Institutions	0.773***	1.068		0.808***	1.074*	
	(0.009)	(0.042)		(0.009)	(0.042)	
Other	0.408***	1.053*	0.118***	0.366***	1.019	0.058***
	(0.015)	(0.028)	(0.182)	(0.016)	(0.027)	(0.162)
Collaborations	0.875	1.306*	0.895*	0.930	1.243	0.925
	(0.005)	(0.180)	(0.062)	(0.005)	(0.175)	(0.056)
Citing parameters						
Company	0.683***	1.149***	0.222***	0.650***	1.189***	0.157***
	(0.014)	(0.037)	(0.011)	(0.014)	(0.038)	(0.009)
Company Collab.	0.907***	0.988	0.104***	0.968***	0.962	0.057***
	(0.007)	(0.087)	(0.108)	(0.008)	(0.085)	(0.094)
Government	0.898***	1.110**	0.261***	0.968***	1.128**	0.194***
	(0.009)	(0.055)	(0.042)	(0.009)	(0.054)	(0.036)
Institutions	0.947***	0.804***		0.925***	0.877***	
	(0.009)	(0.050)		(0.009)	(0.049)	
Other	0.557***	1.703***	0.037***	0.465***	1.593***	0.027***
	(0.019)	(0.055)	(0.013)	(0.020)	(0.051)	(0.012)
Collaborations	0.981***	0.909	1.579***	1.021***	0.917	0.781***
	(0.005)	(0.148)	(0.115)	(0.005)	(0.143)	(0.078)
Decay	0.187***	0.088***	0.058***	0.209***	0.094***	0.054***
	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.005)
Diffusion	0.0002***	0.0003***	2.89E-05*	0.0004***	0.0003***	1.04E-04***
	(4.94E-06)	(1.24E-05)	(0.00002)	(7.30E-06)	(1.41E-05)	(0.00003)
Num. of obs.	1,666,298	469,168	61,573	1,666,298	469,168	61,573

* p<0.10, ** p<0.05, *** p<0.01
Standard errors in parentheses. Governments and institutions combined for patent-article citations due to smaller sample size. Regressions also include controls for the publication year of the citing document and the citing and cited country.

Table 10 – Knowledge flows across organizations, by technology

	Biofuels		Solar		Wind	
	art-art	pat-pat	art-art	pat-pat	art-art	pat-pat
Cited parameters						
Company	0.604***	1.223**	0.811***	1.088***	1.621***	1.049
	(0.019)	(0.089)	(0.009)	(0.030)	(0.080)	(0.065)
Company Collab.	0.777***	0.323	1.102***	0.465***	2.150***	1.353
	(0.012)	(0.426)	(0.006)	(0.130)	(0.044)	(0.221)
Government	0.828***	1.492***	0.828***	0.906***	1.797***	1.149*
	(0.011)	(0.131)	(0.007)	(0.033)	(0.047)	(0.090)
Institutions	0.594***	1.556***	0.859***	0.899**	1.026	1.043
	(0.016)	(0.141)	(0.009)	(0.043)	(0.051)	(0.101)
Other	0.468***	1.057	0.297***	1.033	0.690***	1.081
	(0.025)	(0.092)	(0.018)	(0.028)	(0.046)	(0.066)
Collaborations	0.741***	1.033	0.942***	1.507***	1.742***	N/A
	(0.008)	(0.717)	(0.005)	(0.138)	(0.032)	
Citing parameters						
Company	0.815***	1.754***	0.772***	1.703***	0.315***	0.478***
	(0.025)	(0.203)	(0.015)	(0.076)	(0.041)	(0.022)
Company Collab.	0.873***	1.566*	0.941***	1.217	0.525***	0.373***
	(0.013)	(0.337)	(0.007)	(0.172)	(0.031)	(0.077)
Government	0.775***	0.933	0.999	1.362	0.508***	0.688***
	(0.015)	(0.183)	(0.010)	(0.078)	(0.066)	(0.059)
Institutions	0.764***	0.608***	1.046***	1.456***	1.125**	0.462***
	(0.017)	(0.148)	(0.008)	(0.086)	(0.055)	(0.055)
Other	0.582***	5.022***	0.489***	2.277***	0.643***	0.735***
	(0.030)	(0.575)	(0.024)	(0.102)	(0.047)	(0.033)
Collaborations	0.798***	1.981	1.095***	2.329***	0.951*	N/A
	(0.009)	(1.149)	(0.005)	(0.249)	(0.025)	
Decay	0.147***	0.062***	0.219***	0.111***	0.131***	0.079***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Diffusion	0.0003***	0.0005***	0.0003***	0.0003***	2.6E-05***	0.0020***
	(1.33E-05)	(1.15E-04)	(5.80E-06)	(1.46E-05)	(4.20E-06)	(1.73E-04)
Num. of obs.	604,135	45,104	666,779	291,285	395,384	132,779

* p<0.10, ** p<0.05, *** p<0.01
Standard errors in parentheses. Results excluding self-citations. Regressions also include controls for the publication year of the citing document and the citing and cited country.

ON-LINE SUPPLEMENT: APPENDIX TABLES AND FIGURES

Appendix Table A1 – Keyword Searches for Energy Articles

Biofuels

#11	TS = ("biomass" NEAR/5 "electricit*" OR "biomass fuel*" OR "biomass heat*" OR "biomass energy" OR "Bio feedstock*" OR "biofeedstock*" OR "Hydrotreated vegetable oil*" or "lignocellulosic biomass*" OR "cellulosic ethanol*" or "biomass to liquid*" OR "bio synthetic gas*" OR "algae-based fuel*" OR "landfill gas*" or "Biohydrogen production*" or "Biological hydrogen production*" or "bio energy" or "bioenergy" or "biofuel*" or "bio fuel*" or "biodiesel*" or "bio diesel*" or "biogas*" or "bio gas*" OR "Bio syngas*" or "bio oil" or "bio ethanol*" or "bioethanol*" OR "fuel ethanol*" OR "Biomethanol*" OR "bio methanol*") NOT TS = ("co-combust*" or "cocombust*" or "co-fir*" or "cofir*" or "multi-combust*" or "multicombust*" or "multi-fir*" or "multifir*" or "fuel cell*" or "biofuel cell*") Refined by: Web of Science Core Collection Categories=(ENERGY FUELS OR SPECTROSCOPY OR BIOTECHNOLOGY APPLIED MICROBIOLOGY OR ENTOMOLOGY OR ENGINEERING CHEMICAL OR ENVIRONMENTAL SCIENCES OR POLYMER SCIENCE OR AGRICULTURAL ENGINEERING OR ENGINEERING ENVIRONMENTAL OR GEOSCIENCES MULTIDISCIPLINARY OR CHEMISTRY MULTIDISCIPLINARY OR TRANSPORTATION SCIENCE TECHNOLOGY OR FOOD SCIENCE TECHNOLOGY OR CHEMISTRY PHYSICAL OR CHEMISTRY APPLIED OR GENETICS HEREDITY OR BIOCHEMISTRY MOLECULAR BIOLOGY OR BIOLOGY OR WATER RESOURCES OR THERMODYNAMICS OR CHEMISTRY ORGANIC OR AGRONOMY OR PHYSICS ATOMIC MOLECULAR CHEMICAL OR GEOCHEMISTRY GEOPHYSICS OR PLANT SCIENCES OR ENGINEERING MECHANICAL OR CHEMISTRY ANALYTICAL OR MULTIDISCIPLINARY SCIENCES OR METEOROLOGY ATMOSPHERIC SCIENCES OR MATERIALS SCIENCE BIOMATERIALS OR AGRICULTURE MULTIDISCIPLINARY OR DEVELOPMENTAL BIOLOGY OR MICROBIOLOGY OR ECOLOGY OR MECHANICS OR ENGINEERING INDUSTRIAL OR FORESTRY OR HORTICULTURE OR BIOCHEMICAL RESEARCH METHODS OR NANOSCIENCE NANOTECHNOLOGY OR ENGINEERING MULTIDISCIPLINARY OR SOIL SCIENCE OR MATERIALS SCIENCE PAPER WOOD OR METALLURGY METALLURGICAL ENGINEERING OR MATERIALS SCIENCE TEXTILES OR ELECTROCHEMISTRY OR ENGINEERING CIVIL OR MATERIALS SCIENCE MULTIDISCIPLINARY) Timespan=1991-2010. Databases=SCI-EXPANDED. Lemmatization=Off

Solar Energy

The solar energy search combines searches for specific types of solar energy (e.g. solar PV) and a more general search strategy:

Solar Thermal Power

#1	(TS = (solar NEAR/2 thermoelectr*) OR TS = (solar NEAR/2 “power plant”) OR TS = (“concentrat* solar” NEAR/2 power) OR TS= (“solar thermal” NEAR/2 (power OR electric*)) OR TS=(parabolic* NEAR/2 trough*) OR TS=((parabolic NEAR/2 dish*) AND solar) OR TS = (stirling NEAR/2 dish*) OR TS=((Fresnel NEAR/2 (reflector* OR lens*)) AND solar)) NOT (TS = (cell* OR photovoltaic* OR PV) OR TS = (hydrogen NEAR/1 (generat* or product*)) OR TS = (battery OR batteries) OR TS = (storage OR store OR storing)) Refined by: [excluding] Web of Science Core Collection Categories= (ENGINEERING AEROSPACE OR ASTRONOMY ASTROPHYSICS) Timespan=1991-2010. Databases=SCI-EXPANDED. Lemmatization=Off
#2	TS=(solar NEAR/2 tower) NOT (TS = (cell* OR photovoltaic* OR PV) OR TS = (hydrogen NEAR/1 (generat* or product*)) OR TS = (battery OR batteries) OR TS = (storage OR store OR storing)) Refined by: [excluding] Web of Science Core Collection Categories= (ASTRONOMY ASTROPHYSICS OR NUCLEAR SCIENCE TECHNOLOGY OR METEOROLOGY ATMOSPHERIC SCIENCES) Timespan=1991-2010. Databases=SCI-EXPANDED. Lemmatization=Off
Solar Photovoltaic	
#3	TS = ("photovoltaic energ*" OR "solar cell*" OR "photovoltaic power *" OR "photovoltaic cell*" OR "photovoltaic solar energy*") NOT (TS = (hydrogen NEAR/1 (generat* or product*)) OR TS = (battery OR batteries) OR TS = (storage OR store OR storing))

Solar General	
#4	TS = (“solar panel*” OR “solar array*” OR “solar resource*” OR “solar potential” OR “solar energy” OR “solar collector*”) NOT (#5 OR #8 OR #9 OR TS = (hydrogen NEAR/1 (generat* or product*)) OR TS = (battery OR batteries) OR TS = (storage OR store OR storing)) Refined by: Web of Science Core Collection Categories= (AUTOMATION CONTROL SYSTEMS OR CHEMISTRY ANALYTICAL OR CHEMISTRY INORGANIC NUCLEAR OR CHEMISTRY MULTIDISCIPLINARY OR CHEMISTRY ORGANIC OR CHEMISTRY PHYSICAL OR CONSTRUCTION BUILDING TECHNOLOGY OR ELECTROCHEMISTRY OR ENERGY FUELS OR ENGINEERING CIVIL OR ENGINEERING ELECTRICAL ELECTRONIC OR ENGINEERING MULTIDISCIPLINARY OR ENVIRONMENTAL SCIENCES OR IMAGING SCIENCE PHOTOGRAPHIC TECHNOLOGY OR MATERIALS SCIENCE CERAMICS OR MATERIALS SCIENCE COATINGS FILMS OR MATERIALS SCIENCE MULTIDISCIPLINARY OR MECHANICS OR METALLURGY METALLURGICAL ENGINEERING OR MINING MINERAL PROCESSING OR NANOSCIENCE NANOTECHNOLOGY OR OPTICS OR PHYSICS APPLIED OR PHYSICS CONDENSED MATTER OR PHYSICS NUCLEAR OR POLYMER SCIENCE OR THERMODYNAMICS OR WATER RESOURCES) AND [excluding] Web of Science Core Collection Categories=(METEOROLOGY ATMOSPHERIC SCIENCES OR ENGINEERING AEROSPACE OR ASTRONOMY ASTROPHYSICS) Timespan=1991-2010. Databases=SCI-EXPANDED. Lemmatization=Off

Wind Energy

#1	TS = ("wind power" OR "wind energy" OR "wind turbine*" OR "wind farm*" OR "wind park*" OR "wind plant*") NOT TS = (battery OR batteries OR storage OR store OR storing OR "hydrogen production*" OR "wind" NEAR "hydrogen" OR "grid integration*" OR "load management" OR "offshore" NEAR/5 ("connect*" OR "link*" OR "electric*" OR "grid*")) Refined by: Web of Science Core Collection Categories=(ENERGY FUELS OR MATERIALS SCIENCE COMPOSITES OR ENGINEERING ELECTRICAL ELECTRONIC OR ORNITHOLOGY OR ENGINEERING MECHANICAL OR ENVIRONMENTAL SCIENCES OR COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE OR MECHANICS OR MATERIALS SCIENCE CHARACTERIZATION TESTING OR ENGINEERING CIVIL OR PHYSICS MULTIDISCIPLINARY OR THERMODYNAMICS OR STATISTICS PROBABILITY OR MATHEMATICS INTERDISCIPLINARY APPLICATIONS OR METEOROLOGY ATMOSPHERIC SCIENCES OR ENGINEERING MARINE OR ENGINEERING MULTIDISCIPLINARY OR ECOLOGY OR METALLURGY METALLURGICAL ENGINEERING OR AUTOMATION CONTROL SYSTEMS OR INSTRUMENTS INSTRUMENTATION OR MATERIALS SCIENCE MULTIDISCIPLINARY OR MULTIDISCIPLINARY SCIENCES OR BIOLOGY OR PHYSICS APPLIED OR COMPUTER SCIENCE THEORY METHODS OR ENGINEERING AEROSPACE OR CONSTRUCTION BUILDING TECHNOLOGY OR REMOTE SENSING OR ENGINEERING OCEAN OR OPERATIONS RESEARCH MANAGEMENT SCIENCE OR ACOUSTICS OR COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS OR MARINE FRESHWATER BIOLOGY OR ENGINEERING INDUSTRIAL OR ZOOLOGY OR PHYSICS MATHEMATICAL OR MATHEMATICS APPLIED) AND [excluding] Web of Science Core Collection Categories= (ASTRONOMY ASTROPHYSICS OR GEOSCIENCES MULTIDISCIPLINARY) Databases=SCI-EXPANDED Timespan=1991-2010 Lemmatization=Off

Appendix Table A2 – IPC codes for Energy Patents

Biofuels	
C12P 7/06-14	BIOCHEMISTRY; BEER; SPIRITS; WINE; VINEGAR; MICROBIOLOGY; ENZYMOLOGY; MUTATION OR GENETIC ENGINEERING/ FERMENTATION OR ENZYME-USING PROCESSES TO SYNTHESISE A DESIRED CHEMICAL COMPOUND OR COMPOSITION OR TO SEPARATE OPTICAL ISOMERS FROM A RACEMIC MIXTURE/ Preparation of oxygen-containing organic compounds/containing a hydroxy group/acyclic/ethanol, i.e. non-beverage
NOT C12P 7/12	BIOCHEMISTRY; BEER; SPIRITS; WINE; VINEGAR; MICROBIOLOGY; ENZYMOLOGY; MUTATION OR GENETIC ENGINEERING/ FERMENTATION OR ENZYME-USING PROCESSES TO SYNTHESISE A DESIRED CHEMICAL COMPOUND OR COMPOSITION OR TO SEPARATE OPTICAL ISOMERS FROM A RACEMIC MIXTURE/ Preparation of oxygen-containing organic compounds/containing a hydroxy group/acyclic/ ethanol, i.e. non-beverage/substrate containing sulfite waste liquor or citrus waste
NOT C12C	BIOCHEMISTRY; BEER; SPIRITS; WINE; VINEGAR; MICROBIOLOGY; ENZYMOLOGY; MUTATION OR GENETIC ENGINEERING/BREWING OF BEER
NOT C12G 3	BIOCHEMISTRY; BEER; SPIRITS; WINE; VINEGAR; MICROBIOLOGY; ENZYMOLOGY; MUTATION OR GENETIC ENGINEERING/ WINE; OTHER ALCOHOLIC BEVERAGES; PREPARATION THEREOF/Preparation of other alcoholic beverages
Solar energy	
F03G 6	MACHINES OR ENGINES FOR LIQUIDS; WIND, SPRING, OR WEIGHT MOTORS; PRODUCING MECHANICAL POWER OR A REACTIVE PROPULSIVE THRUST, NOT OTHERWISE PROVIDED FOR/ SPRING, WEIGHT, INERTIA, OR LIKE MOTORS; MECHANICAL-POWER-PRODUCING DEVICES OR MECHANISMS, NOT OTHERWISE PROVIDED FOR OR USING ENERGY SOURCES NOT OTHERWISE PROVIDED FOR /Devices for producing mechanical power from solar energy
F24J 2	MECHANICAL ENGINEERING; LIGHTING; HEATING; WEAPONS; BLASTING/HEATING, RANGES, VENTILATING/PRODUCTION OR USE OF HEAT NOT OTHERWISE PROVIDED FOR/Use of solar heat, e.g. solar heat collectors
H01L 27/142	ELECTRICITY/BASIC ELECTRIC ELEMENTS/ SEMICONDUCTOR DEVICES; ELECTRIC SOLID STATE DEVICES NOT OTHERWISE PROVIDED FOR/ Devices consisting of a plurality of semiconductor or other solid-state components formed in or on a common

	substrate/including semiconductor components specially adapted for rectifying, oscillating, amplifying or switching and having at least one potential-jump barrier or surface barrier; including integrated passive circuit elements with at least one potential-jump barrier or surface barrier/energy conversion devices
H01L 31/04-058	ELECTRICITY/BASIC ELECTRIC ELEMENTS/ SEMICONDUCTOR DEVICES; ELECTRIC SOLID STATE DEVICES NOT OTHERWISE PROVIDED FOR/ Semiconductor devices sensitive to infra-red radiation, light, electromagnetic radiation of shorter wavelength, or corpuscular radiation and specially adapted either for the conversion of the energy of such radiation into electrical energy or for the control of electrical energy by such radiation; Processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof; Details thereof/Adapted as conversion devices/ including a panel or array of photoelectric cells, e.g. solar cells
H02N 6	ELECTRICITY/ GENERATION, CONVERSION, OR DISTRIBUTION OF ELECTRIC POWER/ELECTRIC MACHINES NOT OTHERWISE PROVIDED FOR/ Generators in which light radiation is directly converted into electrical energy

Wind

F03D	MACHINES OR ENGINES FOR LIQUIDS; WIND, SPRING, OR WEIGHT MOTORS; PRODUCING MECHANICAL POWER OR A REACTIVE PROPULSIVE THRUST, NOT OTHERWISE PROVIDED FOR/Wind Motors

Appendix Table 3 – Complete citation regression results, excluding self-citations

	article-to-article		*patent-to-patent*		*patent-to-article*	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Biofuels	1.356	0.006	8.577	0.141	0.987	0.059
Wind	2.540	0.018	3.417	0.038	0.605	0.108
Citing year 1991-92	0.549	0.123	1.354	0.038	N/A[2]	N/A[2]
Citing year 1993-94	0.968	0.057	0.796	0.035	3.210	2.569
Citing year 1995-96	1.037	0.039	1.034	0.039	2.703	1.801
Citing year 1997-98	0.975	0.031	0.926	0.038	0.582	0.771
Citing year 2001-02	1.091	0.026	1.202	0.040	1.039	0.678
Citing year 2003-04	1.296	0.027	1.019	0.036	0.790	0.528
Citing year 2005-06	1.356	0.027	1.367	0.041	5.708	3.218
Citing year 2007-08	1.373	0.027	1.334	0.041	5.527	3.113
Citing year 2009-02	1.204	0.023	1.591	0.045	3.335	1.881
Citing year 2011-12[1]	0.934	0.018	1.066	0.030	2.809	1.582
cited country: Canada	0.741	0.009	0.814	0.020		
cited country: China	0.549	0.005				
cited country: Denmark	0.769	0.011	0.696	0.029		
cited country: other EU	0.632	0.004	0.656	0.016		
cited country: Germany	0.582	0.005	0.545	0.014		
cited country: Japan	0.586	0.004	0.656	0.015		
cited country: other	0.522	0.003	0.706	0.015		
cited country: Spain	0.633	0.007	0.915	0.048		
cited country: Turkey	0.672	0.008				
cited country: UK	0.804	0.006	0.780	0.028		
citing country: Canada	0.999	0.012	0.902	0.020		
citing country: China	1.130	0.007				
citing country: Denmark	0.770	0.016	0.434	0.021		
citing country: other EU	0.861	0.006	0.671	0.016		
citing country: Germany	0.774	0.008	0.572	0.011		
citing country: Japan	0.896	0.008	0.395	0.011		
citing country: other	0.939	0.005	0.620	0.012		
citing country: Spain	1.033	0.011	0.755	0.036		
citing country: Turkey	1.269	0.013				
cited country: UK	1.012	0.010	0.687	0.027		
cited country: non-US					0.332	0.016
citing country: non-US					0.211	0.026

cited org: company	0.728	0.009	1.041	0.027	0.532	0.081
cited org: comp. collaboration	1.005	0.006	0.832	0.109	0.648	0.068
cited org: government	0.858	0.007	1.136	0.037		
cited org: research institute	0.773	0.009	1.068	0.042		
cited org: govt./institute					1.142	0.063
cited org: other	0.408	0.015	1.053	0.028	0.118	0.182
cited org: collaboration	0.875	0.005	1.306	0.180	0.895	0.062
citing org: company	0.683	0.014	1.149	0.037	0.222	0.011
citing org: comp. collaboration	0.907	0.007	0.988	0.087	0.104	0.108
citing org: government	0.898	0.009	1.110	0.055		
citing org: research institute	0.947	0.009	0.804	0.050		
citing org: govt./institute					0.261	0.042
citing org: other	0.557	0.019	1.703	0.055	0.037	0.013
citing org: collaboration	0.981	0.005	0.909	0.148	1.579	0.115
Decay	0.187	0.001	0.088	0.001	0.058	0.005
Diffusion	0.00025	4.94E-06	0.00025	0.000012	0.000029	0.000016
Number of obs	1666298		469168		61573	
R^2	0.9988		0.9907		0.9992	
Adjusted R^2	0.9988		0.9907		0.9992	

1: Citing year 2011 only for publication-publication citations

2: There are no patents from 1991-92 citing articles from this database

Figure A1: Biofuel patents: top 5 countries (biofuels_patents_color.emf)

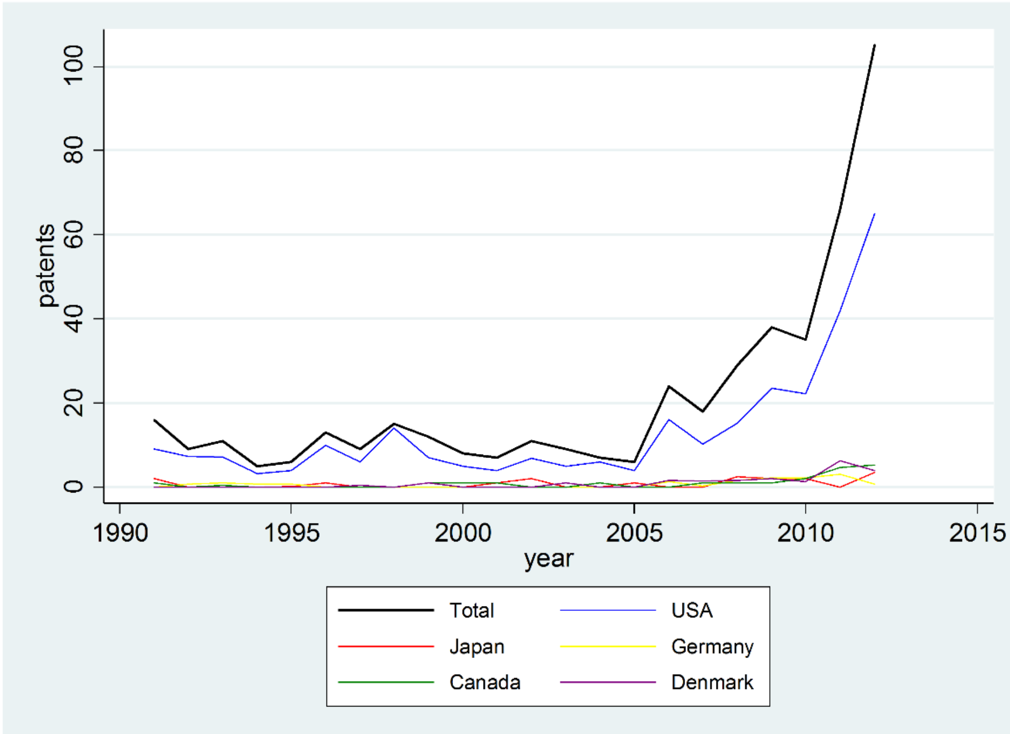

Figure A2: Biofuel articles: top 5 countries (biofuels_publications_color.emf)

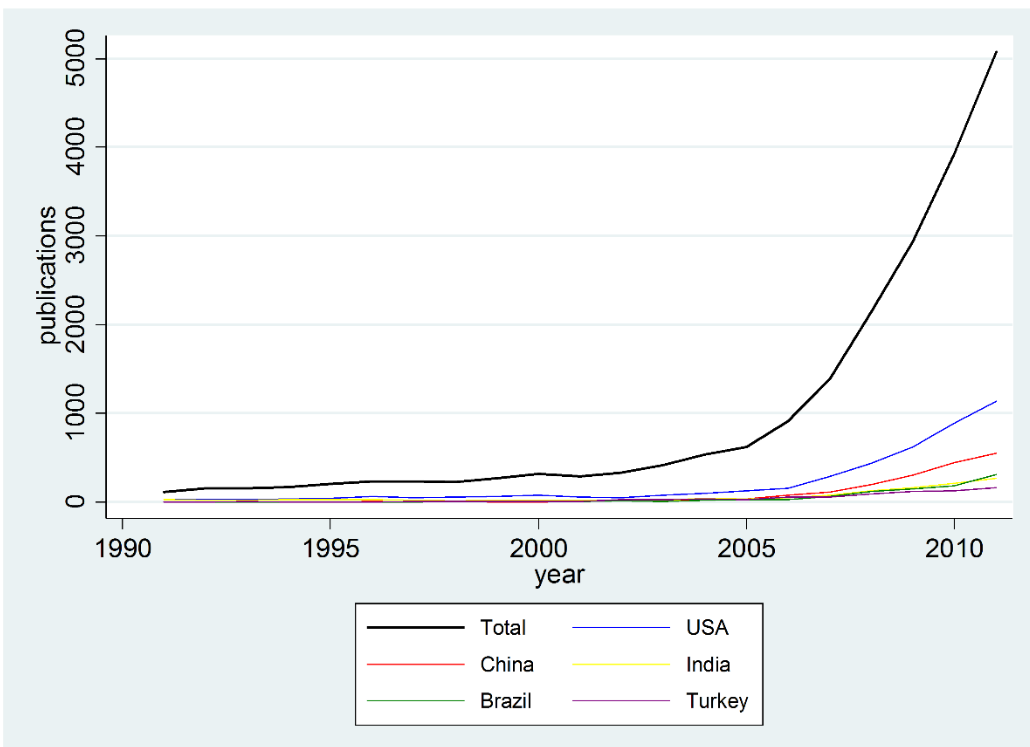

Figure A3: Solar patents: top 5 countries (solar_patents_color.emf)

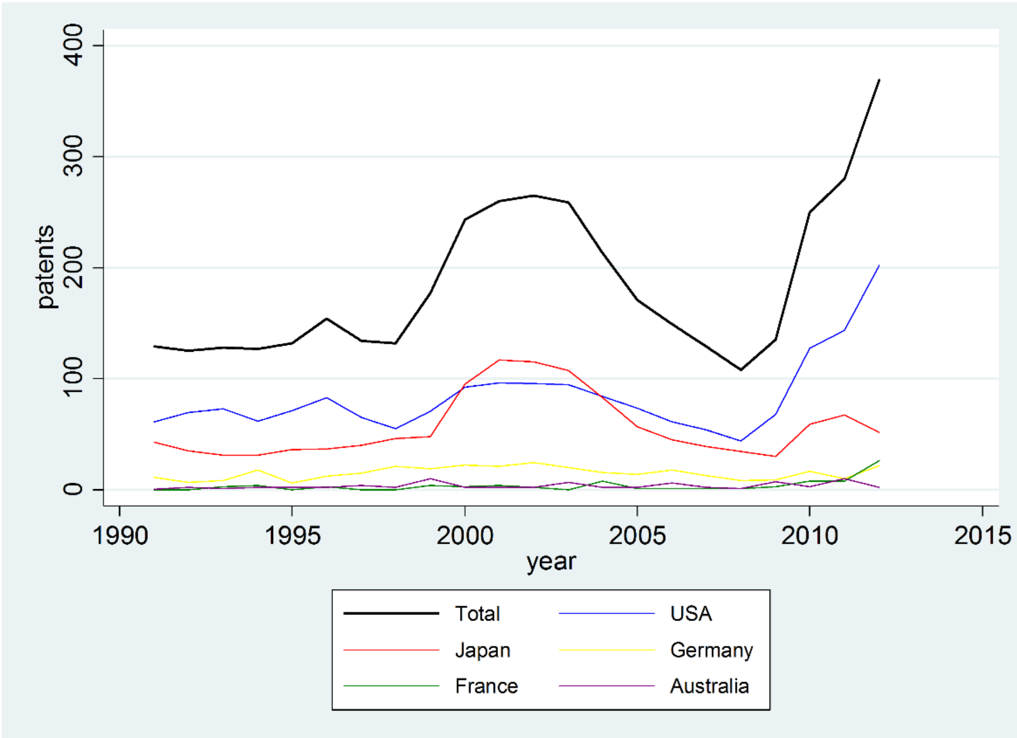

Figure A4: Solar articles: top 5 countries (solar_publications_color.emf)

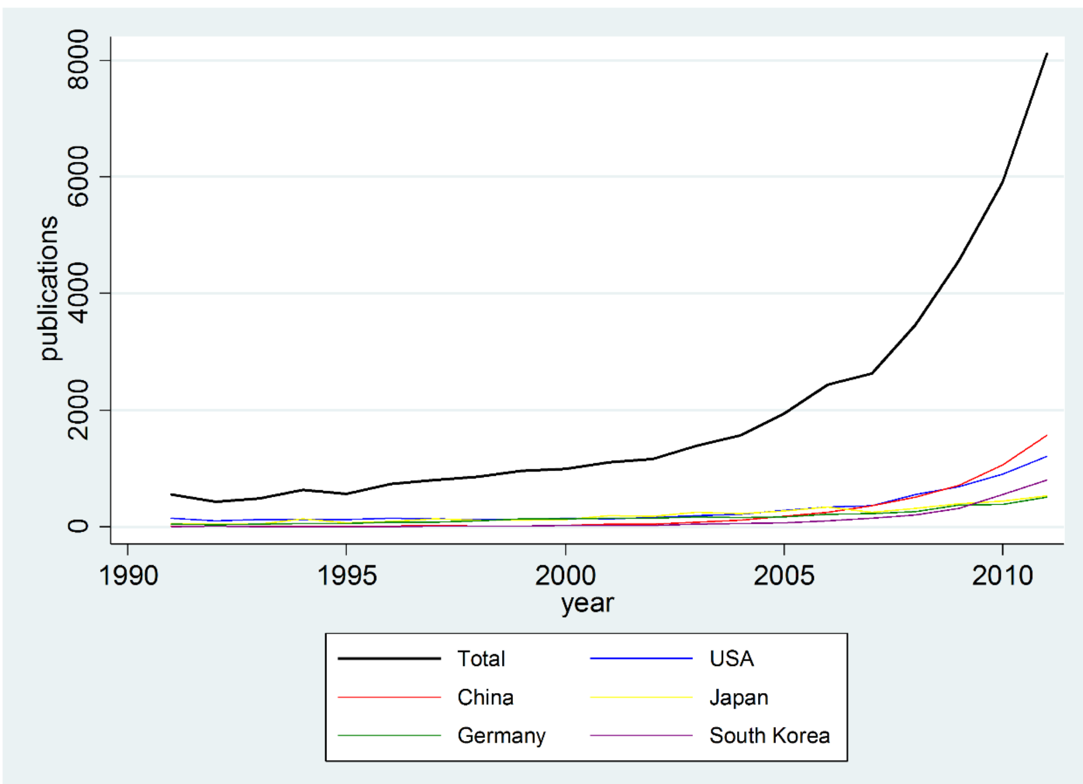

Figure A5: Wind patents: top 5 countries (wind_patents_color.emf)

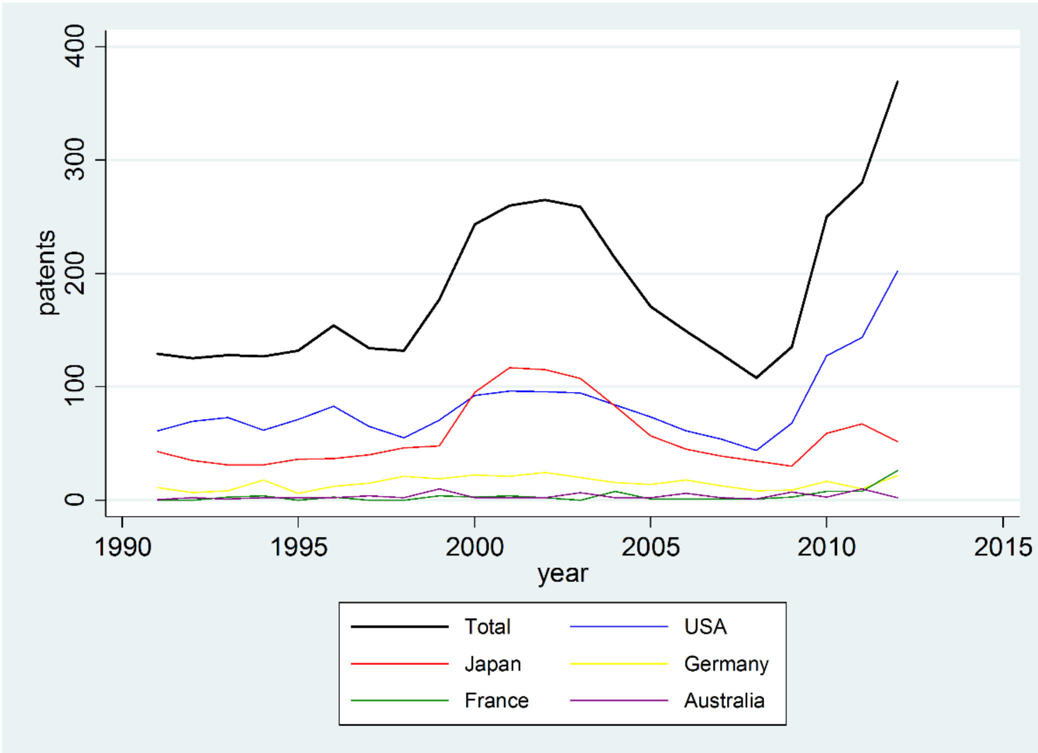

Figure A6: Wind articles: top 5 countries (wind_publications_color.emf)

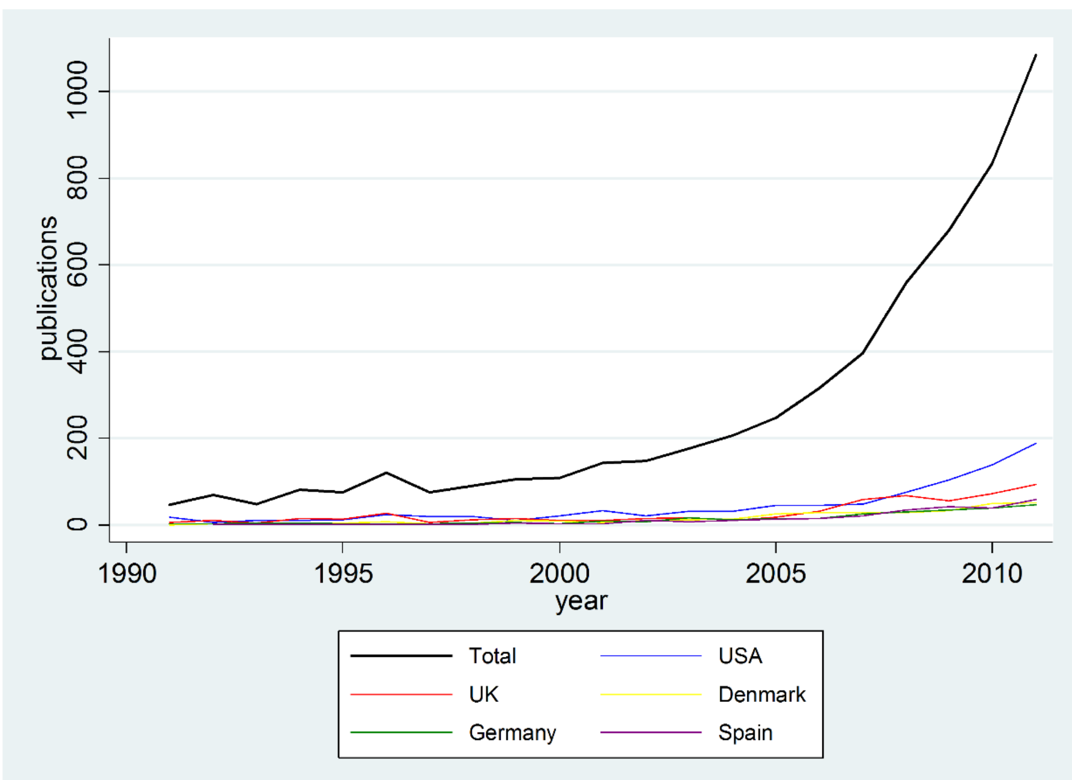

Figure A7: Article author organizations: Biofuels
article_org_piechart_biofuels_COMBINED_country

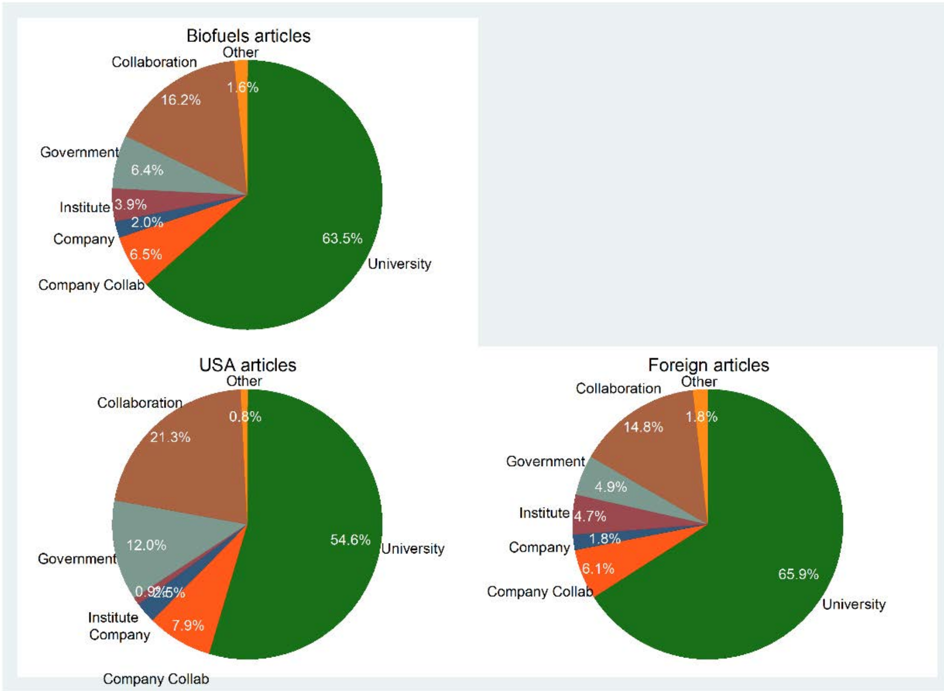

Figure A8: Patent inventor organizations: Biofuels
(patent_org_piechart_biofuels_COMBINED_country.emf)

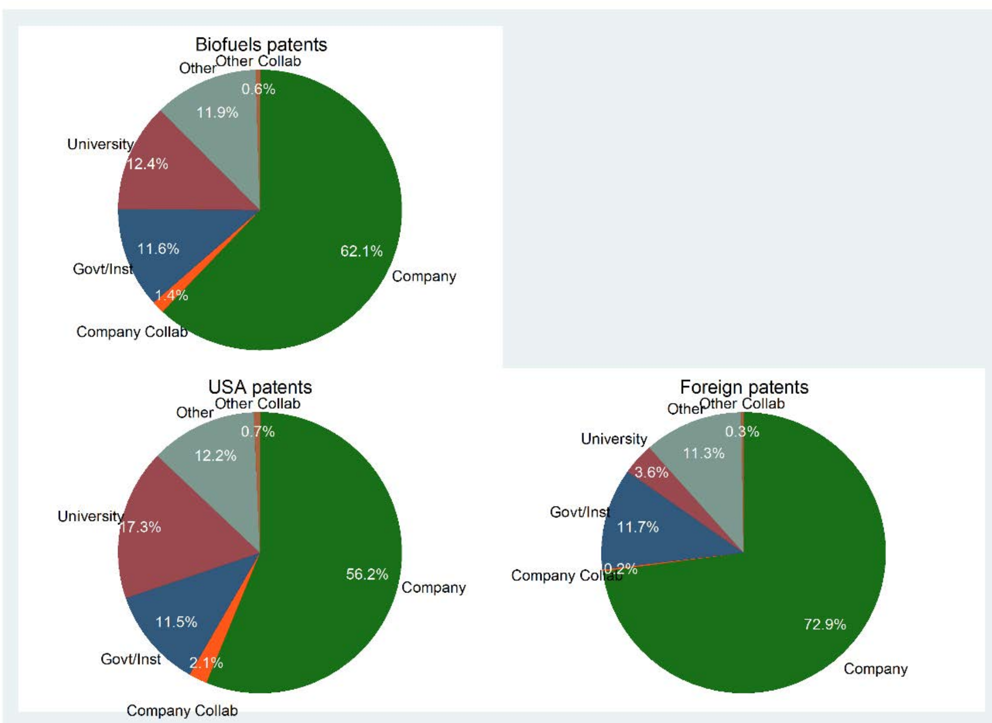

Figure A9: Article author organizations: Solar
(article_org_piechart_solareng_COMBINED_country.emf)

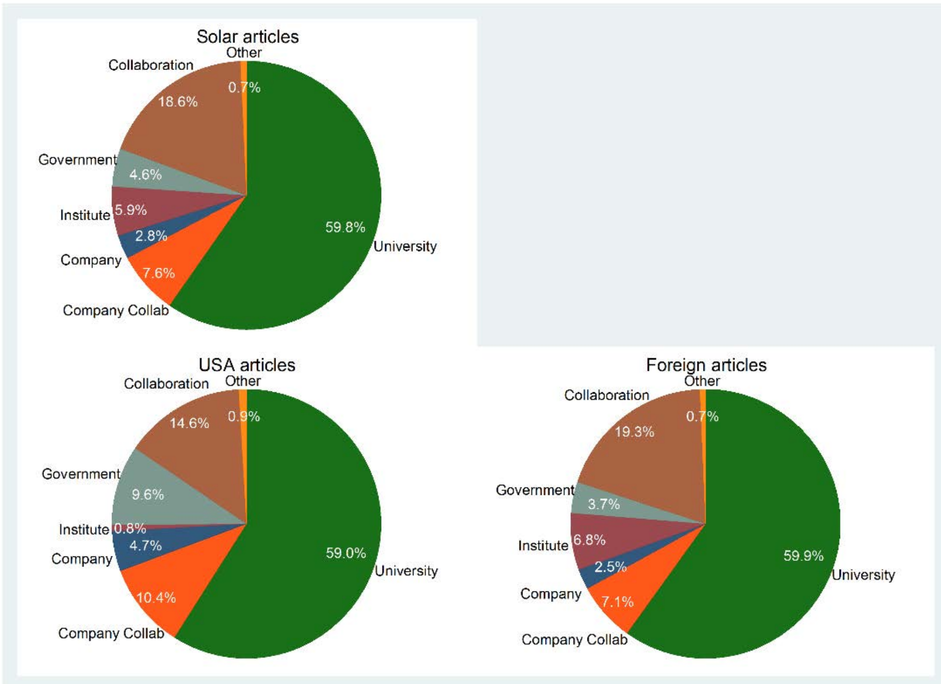

Figure A10: Patent inventor organizations: Solar
(patent_org_piechart_solareng_COMBINED_country.em)

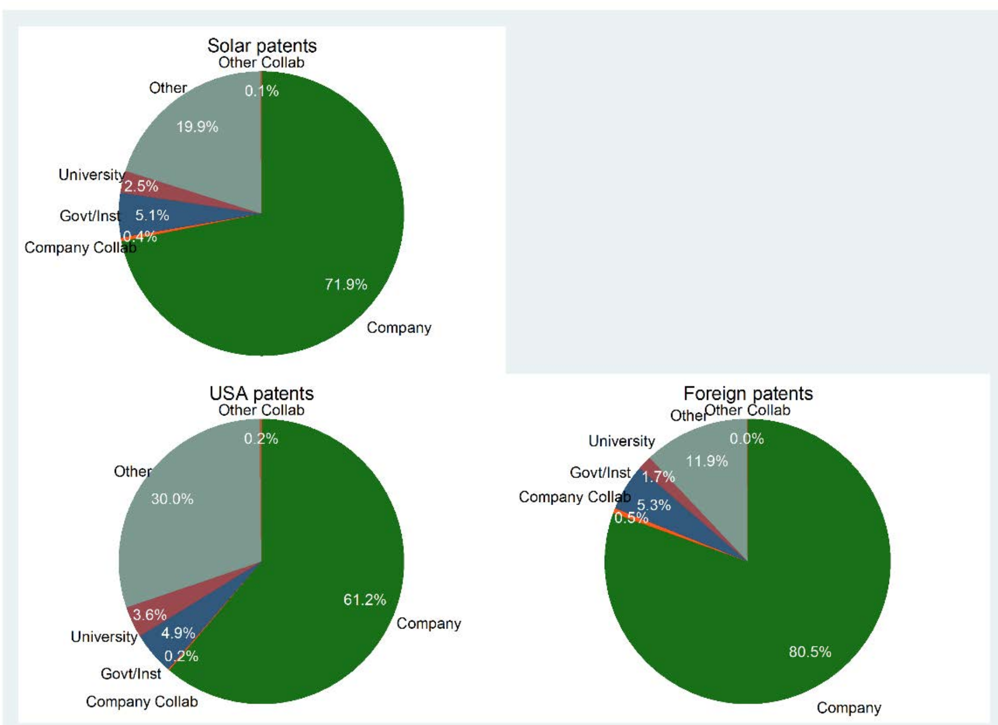

Figure A11: Article author organizations: Wind
article_org_piechart_wind_COMBINED_country

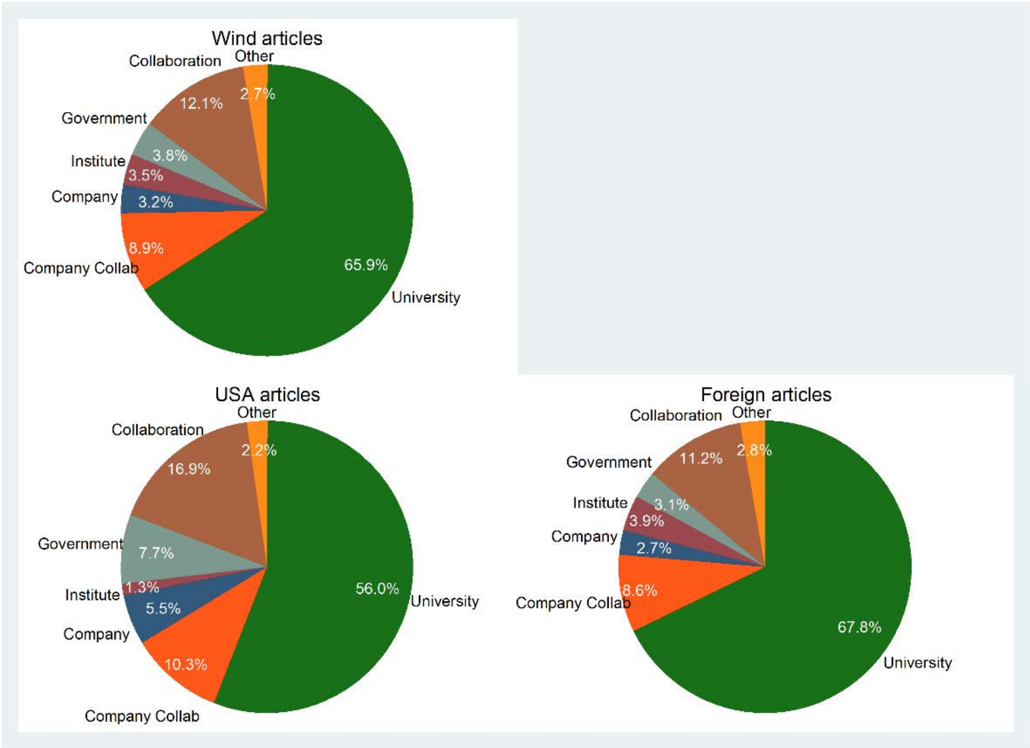

Figure A12: Patent inventor organizations: Wind
(patent_org_piechart_wind_COMBINED_country.emf)

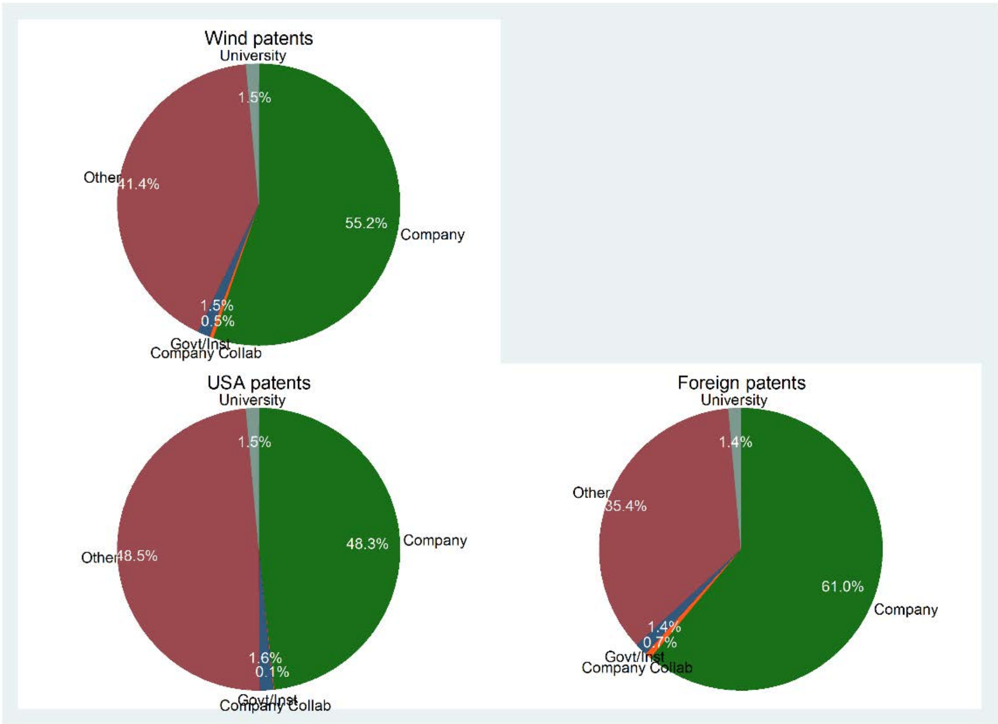

Figure A13: Article author organizations over time: Biofuels
(article_org_piechart_biofuels_COMBINED_years.emf)

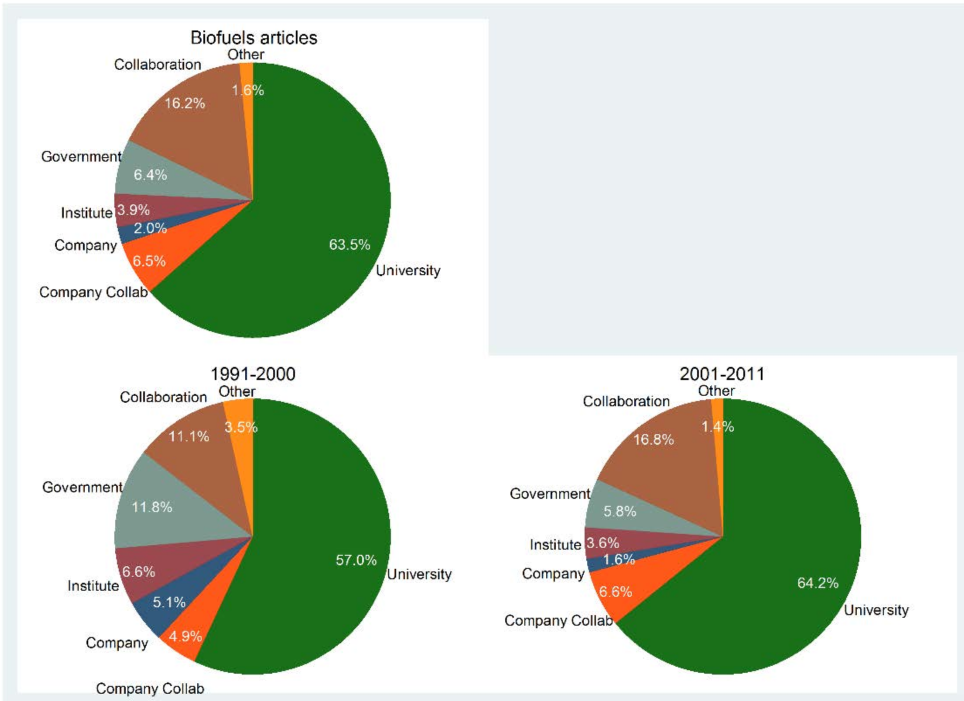

Figure A14: Patent inventor organizations over time: Biofuels
(patent_org_piechart_biofuels_COMBINED_years.emf)

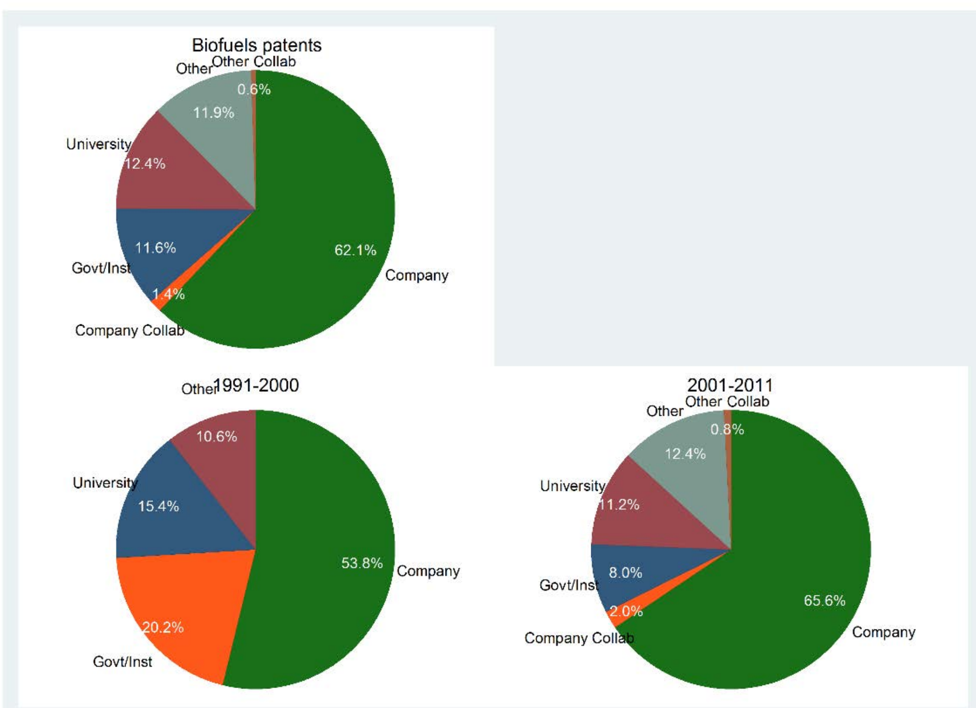

Figure A15: Article author organizations over time: Solar
(article_org_piechart_solareng_COMBINED_years.emf)

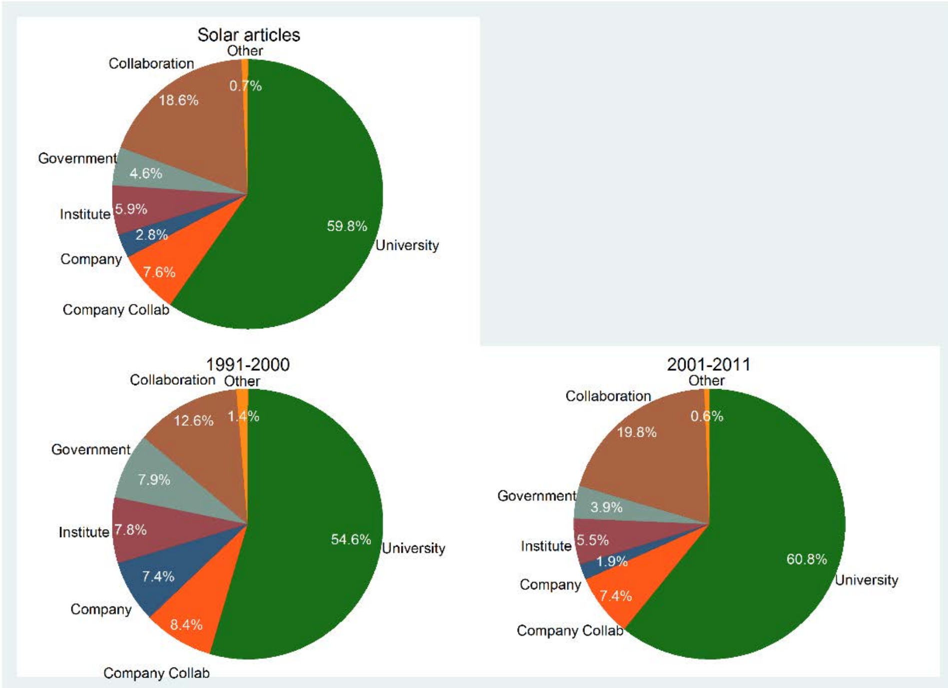

Figure A16: Patent inventor organizations over time: Solar
(patent_org_piechart_solareng_COMBINED_years.emf)

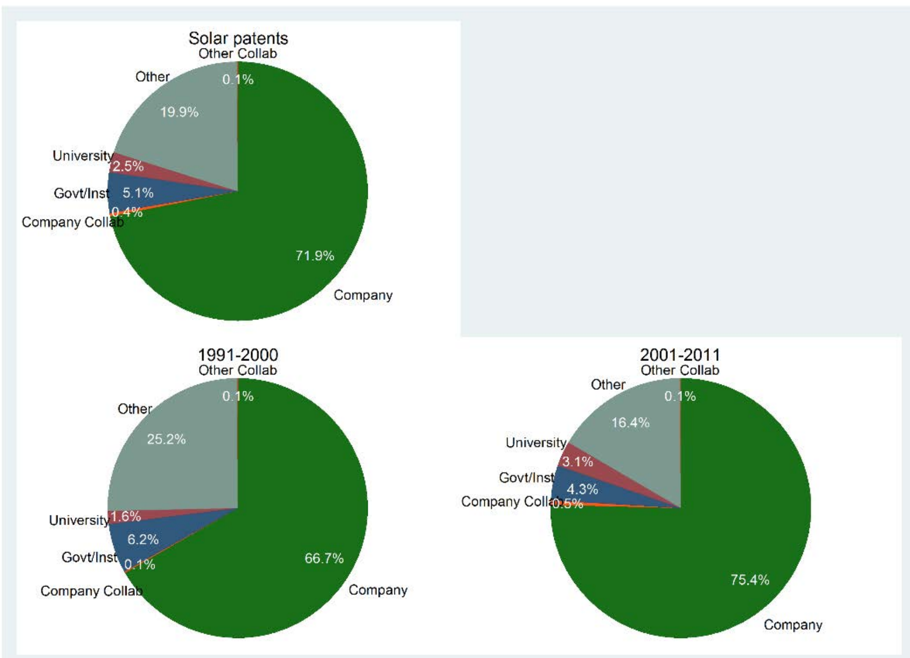

Figure A17: Article author organizations over time: Wind
(article_org_piechart_wind_COMBINED_years.emf)

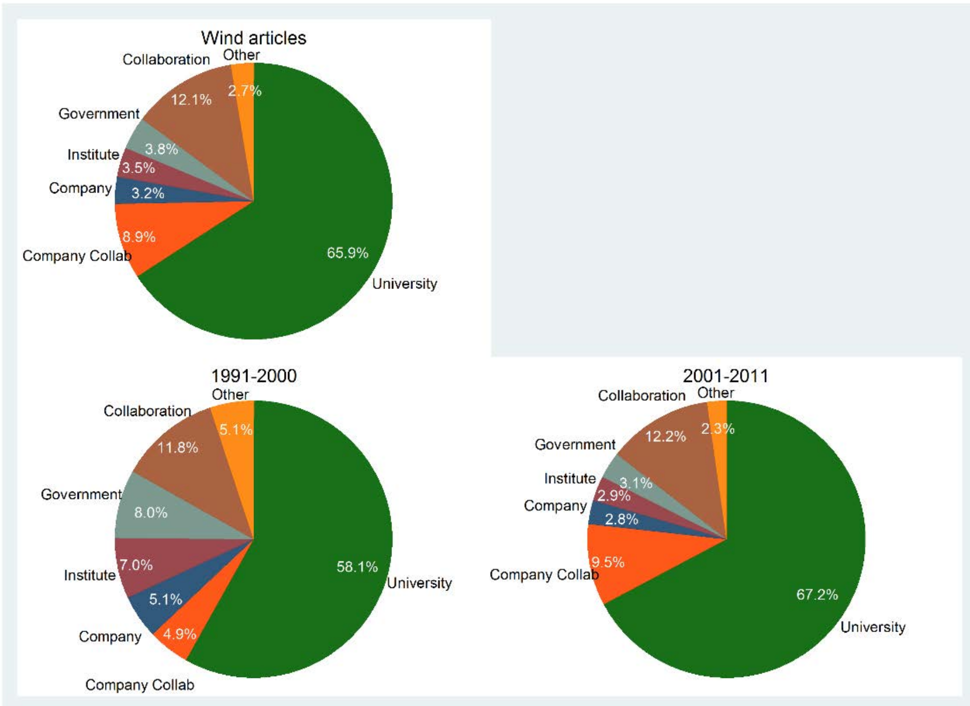

Figure A18: Patent inventor organizations over time:
(patent_org_piechart_wind_COMBINED_years.emf)