



Clean Energy Innovation: Sources of Technical and Commercial Breakthroughs

Thomas D. Perry IV and Mackay Miller
National Renewable Energy Laboratory

Lee Fleming
Harvard Business School

Kenneth Younge
University of Colorado

James Newcomb
National Renewable Energy Laboratory
Current Affiliation: Rocky Mountain Institute

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

Technical Report
NREL/TP-6A20-50624
March 2011

Contract No. DE-AC36-08GO28308

Clean Energy Innovation: Sources of Technical and Commercial Breakthroughs

Thomas D. Perry IV and Mackay Miller
National Renewable Energy Laboratory

Lee Fleming
Harvard Business School

Kenneth Younge
University of Colorado

James Newcomb
National Renewable Energy Laboratory
Current Affiliation: Rocky Mountain Institute

Prepared under Task No. SA10.102C

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

Available electronically at <http://www.osti.gov/bridge>

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information

P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: <mailto:reports@adonis.osti.gov>

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: <http://www.ntis.gov/help/ordermethods.aspx>

Cover Photos: (left to right) PIX 16416, PIX 17423, PIX 16560, PIX 17613, PIX 17436, PIX 17721



Printed on paper containing at least 50% wastepaper, including 10% post consumer waste.

Acknowledgments

We appreciate and acknowledge the support of the Harvard Business School Department of Research, the National Renewable Energy Laboratory, and NSF 0830287 (Creating a Patent Collaboration Network Database to Examine the Social Production of Knowledge). The authors would like to thank R. Lai and A. D'Amour of Harvard Business School and C. Worley of Colorado School of Mines for their assistance with data and analysis and V. Cardona of Heslin Rothenberg Farley & Mesiti P.C., Albany, New York, for sharing their patent search methodology. D. Bilello, T. James, D. Kline, and R. Newmark of NREL and L. Bony of the Rocky Mountain Institute provided helpful comments on the paper. All errors remain ours.

Executive Summary

Low-carbon energy innovation and implementation is essential to combat climate change, promote economic competitiveness, and achieve energy security. Our study of clean energy patenting in the United States is undertaken to elicit fundamental trends and implications that can inform public and private innovation investment, resulting in greater efficiency of research and development programs. Using U.S. patent data and additional patent-relevant data collected from the Internet, we map the landscape of low-carbon energy innovation in the United States since 1975. We isolate 10,603 renewable and 10,442 traditional energy patents and develop a database that characterizes proxy measures for technical and commercial impact, as measured by patent citations and Web presence, respectively. Regression models and multivariate simulations are used to compare the social, institutional, and geographic drivers of breakthrough clean energy innovation. Results indicate statistically significant effects of social, institutional, and geographic variables on technical and commercial impacts of patents and unique innovation trends between different energy technologies. We observe important differences between patent citations and Web presence of licensed and unlicensed patents indicating the potential utility of using screened Web hits as a measure of commercial importance. We offer hypotheses for these revealed differences and suggest a research agenda with which to test these hypotheses. These preliminary findings indicate that leveraging empirical insights to better target research expenditures could augment the speed and scale of innovation and deployment of clean energy technologies.

Table of Contents

| | |
|---|-----------|
| List of Figures | vi |
| List of Tables..... | vi |
| 1 Introduction | 1 |
| 2 Analytical Approach | 4 |
| 2.1 Patent Search..... | 4 |
| 2.2 Patent Interrogation..... | 5 |
| 2.3 Patent Value Definition..... | 5 |
| 2.4 Regressions | 8 |
| 3 Analytical Insights | 10 |
| 3.1 Historical Macro Sources of Clean Energy Inventive Activity | 10 |
| 3.2 Sources of Clean Energy Technical and Commercial Breakthroughs..... | 18 |
| 3.3 Breakthrough Sources..... | 22 |
| 3.4 Predicting Breakthroughs..... | 28 |
| 4 Suggestions for Further Research..... | 30 |
| 5 Conclusions..... | 32 |
| References | 33 |

List of Figures

| | |
|---|----|
| Figure 1. Average patent citations (\bar{y}) per patent by year | 7 |
| Figure 2. Stylized schematic of regression comparison of two related populations (a and b) of patents or Web hits to identify the significance and simulated difference (Δ_{b-a})..... | 8 |
| Figure 3. Annual granted U.S. energy patents (1975–2008) by technology area..... | 10 |
| Figure 4. A comparison of renewable energy patents generated by U.S. and non-U.S. inventors..... | 13 |
| Figure 5. Example collaboration networks for a selected biofuel inventor (dashed circle). | 16 |
| Figure 6. Average collaboration and inventing team size is increasing in energy | 16 |
| Figure 7. Summary network visualizations for each technology area..... | 17 |
| Figure 8. Web hits versus future prior art citations from patents in 1995 (A) and 2005 (B) solar energy patents..... | 20 |
| Figure 9. Signal (A, C) and aggregator (B, D) Web hits versus future prior art citations from patents for licensed and unlicensed National Renewable Energy Laboratory patents | 21 |
| Figure 10. Predictive commercial value of patents using proposed Web hits relationship | 29 |
| Figure 11. An idealized schematic of how future databases might be built and linked in order to understand how science investment influences economic outcomes | 31 |

List of Tables

| | |
|--|----|
| Table 1. Domestic Sources by State for Granted U.S. Energy Patents (1975–2008)..... | 12 |
| Table 2. Market Sources of Patents | 14 |
| Table 3. Regressed and Simulated Determination of Sources for Solar Technical and Commercial Breakthrough Innovation | 24 |
| Table 4. Regressed and Simulated Determination of Sources for Wind Technical and Commercial Breakthrough Innovation | 26 |
| Table 5. Regressed and Simulated Determination of Sources for Biofuel Technical and Commercial Breakthrough Innovation | 27 |

1 Introduction

Development and implementation of low-carbon electricity generation technologies are important for several reasons. To name a few, moving away from fossil-based electricity generation can help increase energy security [1], promote the diversification of energy sources [2], stimulate economic growth [3] and national competitiveness [4], reduce pollution [5], and reduce local water demand [6]. Also, large-scale movement away from practices that create greenhouse gases will help combat climate change.

The scientific consensus on the implications of climate change is serious [7, 8]. Anthropogenic greenhouse gas emissions from energy production and industry sources are causing an “unequivocal ... warming of the climate system” [9]. This warming trend will affect important earth systems that affect global health and prosperity [9, 10]. While the case has been made that we “already possess the fundamental scientific, technical, and industrial know-how to solve the carbon and climate problem for the next half century” [5], the diffusion of low-carbon technologies continues to face significant economic, technical, and political obstacles. Low-carbon technologies can assist in the avoidance and reduction of greenhouse gases and other emissions that have been linked to climate change. Sustained energy technology innovation and market development policies are essential to overcoming these implementation obstacles and are increasingly recognized as a national policy priority for developed and emerging economies alike [11, 12].¹ In this context, the need for rigorous investigation of innovation dynamics of low-carbon technologies is essential to crafting effective innovation and technology transfer policy.

Innovation relies upon a complex ecosystem of fundamental capabilities, an “industrial commons” [13], that serves as a vital source of breakthrough, evolutionary, and incremental innovation [14] and is particularly important for the maturation of new alternative energy products [15]. However, for the past several decades, observers have noted persistent underinvestment in the U.S. energy infrastructure [16], manufacturing [17], and research and development innovation bases [18–20]. While there are indications that some recent research investments are accelerating the development of new technologies [21], several studies indicate insufficient investment [18, 20] to achieve the necessary performance for market adoption [22], to support new technology development [23], and to support the U.S. leadership position in the global economy [24]. For all these reasons, current economic leaders recently called for a massive increase in public renewable energy research and development spending [23].

In recent years, corporate research, development, and deployment dollars have also been important for advancement of the renewable energy economy [25]. However, the recession over the last two years has further hampered domestic corporate research investments [26]. Simultaneously, massive investment in industrial innovative capacity is occurring in other countries, with a specific focus on renewable energy innovation [16, 27]. Because of global competition for market share in alternative energy, innovation in

¹ Useful resources to determine national and regional policies are <http://www.iea.org/textbase/pm/?mode=re> and <http://www.dsireusa.org/>.

energy will be a major contributor to national economic competitiveness in the coming decades. For the United States to establish and maintain a leadership position in the energy economy, new investments are needed to effectively and efficiently stimulate, commercialize, and deploy new energy innovations within our borders and abroad. If this does not occur, the United States runs a risk of falling short of its ability to effectively compete and lead in the global marketplace.

The potential financial opportunities and the opportunity to lead in the global economy have elicited a surge in patenting and investment in clean energy technologies. A leading barometer of inventive activity, the Clean Energy Patent Growth Index [28], shows continuing increases in clean energy patenting activity, yet we currently lack an empirically grounded understanding of how these technologies are invented and enter the market. Policy prescriptions that are a mixture of technology and market stimulation activities may accelerate invention in one industry yet might be ineffectual or even counterproductive in another. For example, if breakthroughs have historically arisen from understanding basic science in one field, and more of these types of innovations are required, then the optimal allocation of funds would be to support early-stage research at universities, national laboratories, or some other similar organization that has previously performed this type of work. On the other hand, if the technology is mature and development of breakthroughs needs to be focused on commercialization and the market, then applied research in government laboratories (most national research facilities increasingly perform a mixture of basic and applied research) or market development activities targeted at private firms might be preferred. Alternatively, fundamental science and understanding may play a far less important role in the development of new technologies. This could manifest if independent inventors were the sources of technical breakthroughs [29] because commercially important innovations could be made without sophisticated experimental facilities (e.g., solar thermal versus photovoltaics). In this case, facilitating commercial deployment of technologies may only require helping functional breakthroughs get to market by providing small business and entrepreneurship loans, incentives for private investment, or other prescriptions to facilitate robust markets for technology [30, 31].

Managing a portfolio of research, development, and deployment investments in the constantly and rapidly evolving energy landscape is a great challenge to public and private research managers. The ultimate aim of this paper is to suggest a line of inquiry to inform research investment decisions across all clean energy technologies and the whole market, ultimately allowing more efficient and effective technology and market investments. This report and its recommendations are preliminary and dependent on the new techniques developed as a part of this work; we expect the scope and specificity of recommendations to continue to evolve as our methodology improves.

This analysis began by asking the question, where do breakthrough innovations in energy originate? While certainly omitting some other potential measures of the advancement of knowledge, such as publications, this investigation uses the patent record as a proxy for innovation [32] and analyzes the sources of breakthrough patents using cutting-edge analytical techniques. This analysis begins by determining patenting trends in biofuel, geothermal, hydro, solar, and wind technologies and continues by mapping the

geographical distribution of patenting activity and illustrating the social networks of each field from co-authorship relationships. While observing gross innovation trends in clean technologies, we also focus our analysis on the source of commercial and technical breakthrough innovations. In particular, it is important to focus on these valuable innovations as unique from bulk patenting trends because most patents are useless or of moderate value [33], as opposed to breakthroughs that significantly change the landscape of technological or business practice. The impact of inventions is then measured in each technology field through two proxy measurements of each patent, technical impact (as measured by future citations) [34], and societal and commercial influence (as quantified by a novel measure of Web presence introduced in this manuscript). These dependent variables enable estimations of the impact of social, institutional, and geographical influences upon the processes of invention for each technology. In the final section of this paper, we provide hypotheses to explain these differences along with a research agenda that we hope can inform policy suggestions.

2 Analytical Approach

2.1 Patent Search

Lai et al. [35] developed a disambiguated database of granted U.S. patents (1975–2008). Using a software text search tool (available at the Harvard Dataverse Network Patent Project [35]) designed to interrogate the U.S. Patent and Trademark (USPTO) database (<http://patft.uspto.gov/>), patent abstracts were searched using energy keyword search terms [18, 20, 36]. We focused our search on technology areas for which search ontologies were already developed and therefore omit other possible areas of important carbon-reduction innovations, such as carbon capture and sequestration (CCS). We experimented with searching only patent titles and the entire patent body for our keywords and found that searching abstracts provided the best mixture of inclusion and accuracy. The resultant lists were cleaned using text analysis functions and by hand to ensure their accuracy. Patents were organized and analyzed by application date to best position the patent within the appropriate inventive context [37] because there is a variable and increasing time lag between patent application and issuance [38].

Our keyword-based search of energy patent titles and abstracts, while effective at identifying patents related to energy technologies, favors patents focused on later-stage innovations that include mention of the production of energy or fuels. For example, when comparing patents that the U.S. Department of Energy (DOE) funded (which presumably are parts of the energy knowledge space) with our keyword-search-driven energy patent list, we observe an overlap of less than 3%. Presumably, this is because the DOE funds very early-stage research (such as algal strains or materials chemistry) that does not lend itself solely to energy generation or fuel applications and may not uniformly include applications to energy in the language of their patent applications. However, while limited to later-stage technologies, our energy list is accurate and includes only energy patents. The accuracy of the keyword search methodology was robust (average of 3.0% initial patent results were determined to not be energy patents after manual inspection) except for nuclear (20.7%) and wind (34.6%) technology areas. The errors in nuclear and wind technology areas were by-in-large a result of patents that included homographic keywords (e.g., “nuclear” in reference to genetic techniques and “turbine” in reference to aeronautic applications, respectively). However, we acknowledge that we are likely omitting many earlier-stage energy technologies because of the keyword selection. We intend to develop a more sophisticated ontology that will combine a more inclusive keyword set while assuring accuracy by using patent classification codes in combination.

Because energy technologies draw heavily upon preceding innovations (e.g., solar draws upon semiconductor manufacturing), this improved ontology will allow interesting investigations of knowledge flow [39] and the organic development of new technology areas [37]. Although this investigation would be interesting, it is beyond the scope of this manuscript. All analyses include only energy patents and not other patents that may have been invented by the included authors. This artificial boundary around the energy patents creates rigid and deterministic barriers for our investigation and limits the determination of the role of the entire inventive career and collaborations therein but is necessary for

this first exploratory work. Follow-on work will provide more comprehensive investigation of the historic social variables involved in energy innovation.

2.2 Patent Interrogation

Three independent variables were developed to provide insight into the structural origins of inventive activity: “Government Interest,” “Assignee Organization Type,” and “References to Science.” Government interest patents are those that the government retains a right to, usually because they funded the work. Patents with government interests were identified using a USPTO search tool [35] to search the government interest field for identifying terms such as “defense” or “energy.” The search terms could not be as rigid as “Department of Defense” because often the patents cited the “Dept. of Defense,” so search terms were developed to be inclusionary. Assignee organization type defines what kind of organization patented the invention. Text search algorithms were developed to search the patent assignee field for terms (e.g., “university,” “corporation,” and “LLC”) that identified the organizational type of the inventing institution. Assignees were classified into five types: Private, University, Government, Independent (i.e., no assignee), and Other. The results of the automated text classification system were then manually reviewed for accuracy. References to science were identified if at least one other non-patent reference in the “Other References” field was coded as referencing science (these references are mostly to scientific publications [40]). While the variable for references to science was included as a covariate control in our model, its influence was not simulated for this investigation.

Three independent variables were developed to provide insight into the social origins of inventive activity: “Team Size,” “Mentorship,” and “Dispersion.” Team size counted the number of inventors listed on the patent application. Patents with two or more inventors were classified as team patents. This variable was included in our simulation results. Mentorship compared the time-in-field of inventors on a patent. If a difference of 10 years or greater existed between any two inventors, the patent was classified as arising from a mentorship relationship. Time-in-field was derived by subtracting the application year of each inventor’s first patent from the application year of the patent in question. Dispersion measured the maximum distance between any two inventors on a single patent, with a distance greater than 100 miles classified as dispersed. Distance between inventors was calculated using the zip code values corresponding to the Inventor City field and so returned values only for patents with two or more U.S. inventors. Mentorship and dispersion variables were included as covariate controls but were not simulated for this manuscript. We also performed qualitative analysis of U.S. inventor patent co-authorship with collaborators in other countries for work unrelated to this manuscript and demonstrated the ability to map collaboration between the United States and other countries. Further research should include work to understand the impact of non-U.S. innovations on domestic innovation, both from the perspective of contribution to the collective knowledge as well as more direct individual collaboration effects.

2.3 Patent Value Definition

2.3.1 Technical Importance: Citation Analytics

Citations have a long history in the bibliometric literature and the first application to patents identified was by Trajtenberg [34]. Heretofore, patent citations are the dominant

means by which differences in patent quality were accessed, especially in large datasets where in-depth qualitative evaluation of individual patents is not possible. Citations to prior patents are submitted as part of a new patent application in order to establish “prior art.” The applicant is required to disclose the prior art to the examiner and must then establish how their invention goes beyond that prior art. The examiner often adds citations they are aware of, such that half of the citations come, on average, from the examiner [41]. While it is not incumbent upon the applicant or the examiner to be exhaustive in their citation inclusion, important patents do on average receive more citations [34]. Similar to publications [42], inventors (authors) often cite their own work in part to show the evolution of knowledge but also, perhaps, to influence their citation relevance [43, 44] or Eigenfactor [45], which are used as the basis of tenure and other promotion decisions [46]. Therefore, for all analysis herein, patent technology value is defined by citations from future patents less inventor self-citations, henceforth denoted simply as citations (γ). Patents that are more highly cited are thought to provide the basis for future technical advance and often change the bases of competition between firms and industries. Hence, we might consider highly cited patents as technical breakthroughs. Besides measuring the impact of individual patents, citations have also been used to measure status and deference [47], knowledge flow [48], and to disambiguate inventor careers [35]. They have also been shown to correlate weakly with financial value as reported by a survey of patent holders [49].

While very popular, we recognize that patent citations are an imprecise and flawed measure of technological importance and breakthrough [50]. Selecting a like patent set for comparison and evaluation of anomalies is essential because it has been observed that different industries have different citation patterns and norms [42] (e.g., the pharmaceutical industry is a highly collaborative and citing industry). Also, since citations accrue to a patent over a period of several years after it issues, citation counts for recent patents are disadvantaged compared to older patents, an effect known as “truncation” [42], and this is true with energy patents as well (Figure 1). In this figure, traditional energy (10,442 patents) and renewable energy (10,603 patents) patent sets were identified by keyword search; the random baseline includes 10,000 randomly selected patents from the entire U.S. patent dataset.

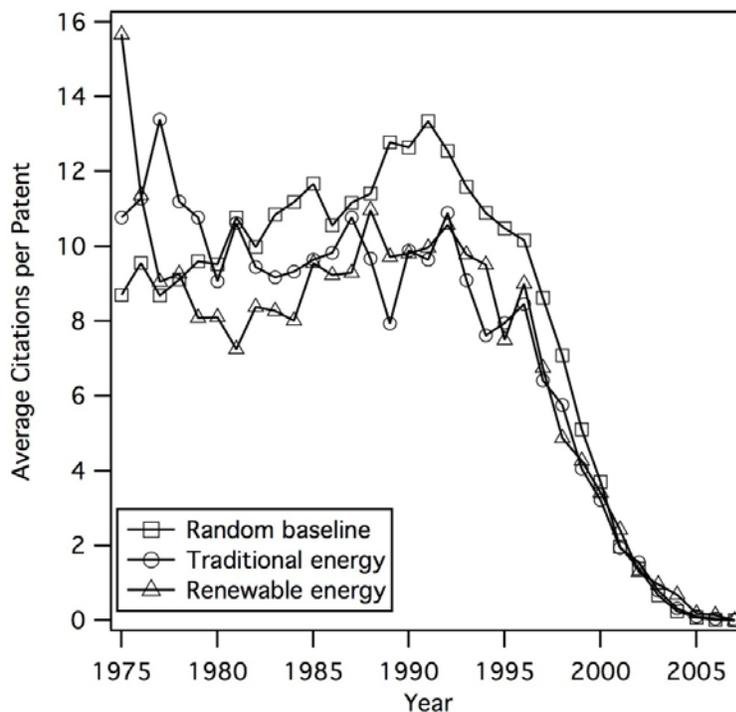


Figure 1. Average patent citations (\bar{y}) per patent by year

Citations also do not strictly measure non-technical or commercial impact, and correlations with these measures likely reflect some underlying but omitted variable. In short, citations measure the impact of invention on future technological trajectories but neglect the commercial and non-technical impact of the technology. However, as a proxy for technological “breakthroughness,” patent citations previously have been used effectively [51] and are currently the best measure available for very large datasets [32]. Given these inherent limitations of citation analysis, this work aims to enrich the measure of patent impact via novel proxies.

2.3.2 Commercial Importance: Web Presence Analytics

We developed and tested a new quantitative measure of Web presence on several patent sets in an attempt to develop a novel proxy measure for quantifying the commercial impact of patents. A count of Web hits (ω) was collected by searching for each patent using the search terms “patent <patent number>,” aggregating the resultant URLs,² and applying an evaluative algorithm to screen the root domains. The evaluative screening algorithm compares URLs to a manually created list of “aggregator” website URLs that essentially replicate information from the USPTO. Those URLs were defined as aggregator URLs (ω_{agg}), while the remaining were defined as potential signal URLs (ω_{sig}). Taken together, these new quantifiable variables for a single or a population of patents can be effectively used to estimate commercial importance.

² In computing, a Uniform Resource Locator (URL) specifies where an identified resource is available and the mechanism for retrieving it. The best known example of the use of URLs is for the addresses of Web pages on the World Wide Web. Accessed December 31, 2010.

2.4 Regressions

To determine the effect of factors of interest (§2.2) on the commercial and technical impact of patents, we estimated long-tailed [33] Poisson models of citation counts and Web hits to clean energy patents. In essence, we determine the statistical likelihood (p value) and difference (Δ_{b-a}) between two populations (a and b) of patents (Figure 2). We use comprehensive pair-wise comparison of like patents to estimate the differences between any two unique factors of interest (e.g., biofuels patents from Colorado versus California or biofuels patents from corporate versus university assignees). An observed and statistically relevant simulated difference indicates that there is an important divergence (positive or negative) between the populations of patents and that the factor of interest is causing some modification in the distribution of the patents' values and may cause or hamper the production of breakthroughs.

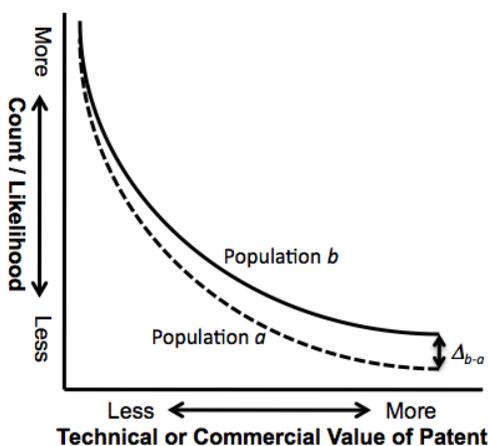


Figure 2. Stylized schematic of regression comparison of two related populations (a and b) of patents or Web hits to identify the significance and simulated difference (Δ_{b-a})

This model uses robust [52] and quasi-maximum-likelihood [53] estimation. Because interactions in non-linear models depend upon the value of other variables in the model [54] as well as the covariance matrix of coefficient estimates in the model [55, 56], we do not interpret marginal effects alone. Instead, we simulate the count at specific levels of the independent variable of interest, holding all other covariates at a reasonable value, such as the mean.³ We assess whether the interaction effects are statistically significantly different by determining the probability that the magnitude of each simulated difference is different from zero [57, 58].

We evaluated the data in several ways in order to ensure the robustness of results. First, only the patents within each field were considered. Second, we ran a full model with all fields that included field-specific interactions. Third, we checked to see if the results held up for a more recent sample of the data. Finally, we ran models with and without controls for yearly fixed effects, such as citation truncation (Figure 1) to remove yearly variation in citation patterns, across all fields. In all models, we coded a focal

³ All discussion of “simulated results” in the paper refers to this alternative method of estimating effect size.

comparison of interest [59] (e.g., between California and Colorado) so that we could examine whether the predicted difference between two different levels was statistically significant. If we saw similar results across multiple models of the data, we flagged this as an opportunity for investigation and performed a qualitative inspection of the patent populations to determine what may be the root causes of the differences.

3 Analytical Insights

This report examines the sources of drivers for renewable energy innovations. In particular, it compares data across clean technology fields to determine the observed historical differences between the fields and the opportunities for tailored innovation policies and strategies that are specific to the energy technology of interest. It describes the revealed sources of energy innovations from a variety of perspectives and then identifies the morphologies of innovation in each technology area that have historically led to breakthrough technical and commercial innovations.

3.1 Historical Macro Sources of Clean Energy Inventive Activity

Starting with the U.S. patent database of over 3.7 million patents, energy patents were identified by keyword search of the USPTO database as belonging to one of nine technology areas: biofuels, coal, geothermal, hydro, natural gas, nuclear, oil, solar (photovoltaics and solar thermal), and wind (Figure 3). Given the low amount of identified patenting in geothermal, hydro, and nuclear technologies, there was little signal in the count models, and we have focused our analysis of renewable energy patents in the wind, biofuels, and solar renewable technology areas.

In response to the 1970s oil crisis, the United States rallied the largest investment in energy in the last several decades [11] that resulted in a peak in energy patenting in the late 1970s (Figure 1 and Figure 3). Since then, funding steadily declined and reached a low point in the mid-1990s. A second surge has recently occurred likely due to the perceived growing economic opportunity and resultant corporate investment, particularly in wind and solar technologies (the very recent fall-off in all technologies is likely an artifact of application and granting delay at the USPTO [37]).

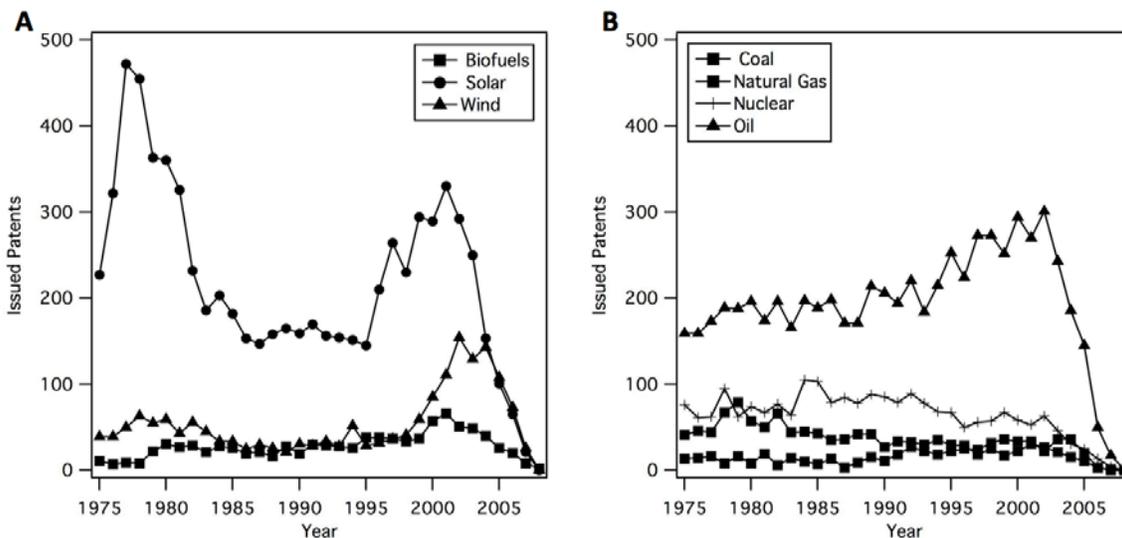


Figure 3. Annual granted U.S. energy patents (1975–2008) by technology area

3.1.1 Geographic Sources

In the United States, we observe increasing inventive activity in the clean energy sector over the last several decades along with increasing geographic concentration (Table 1). California is a dominant geography for energy inventive activity across all technology areas, especially solar, which is likely based on the concentration of the semiconductor and venture capital industries there. High patenting activity is also observed in Texas, Massachusetts, Colorado, and New York. Pennsylvania shows active patenting across all traditional energy technologies, perhaps due to geographic choices in the early oil industry. And not surprisingly, Texas shows strong oil-related patenting activity due to the strength of the industry there. Because of the dominance of California, Massachusetts, New York, and Texas and the high levels of activity in Colorado, we focus our analysis on breakthroughs in those states. In the future, we can extend this analysis for all states—and similarly, any feature of interest that we identified—and run comparisons between them.

Table 1. Domestic Sources by State for Granted U.S. Energy Patents (1975–2008)

| State | Energy Technology | | | | | | | | |
|-------|-------------------|------------|---------|-------|------|-------------|----|-----|-----|
| | Renewable | | | | | Traditional | | | |
| | Biofuels | Geothermal | Hydro | Solar | Wind | Natural | | | Oil |
| | Coal | Gas | Nuclear | | | | | | |
| AK | 0 | 0 | 0 | 4 | 4 | 2 | 1 | 0 | 15 |
| AL | 6 | 0 | 0 | 35 | 6 | 3 | 0 | 6 | 15 |
| AR | 5 | 1 | 0 | 10 | 1 | 1 | 0 | 2 | 4 |
| AZ | 7 | 2 | 4 | 162 | 25 | 1 | 0 | 11 | 45 |
| CA | 65 | 110 | 12 | 1,190 | 218 | 64 | 55 | 252 | 416 |
| CO | 27 | 2 | 2 | 205 | 48 | 31 | 13 | 3 | 72 |
| CT | 11 | 0 | 3 | 95 | 39 | 51 | 15 | 70 | 92 |
| DC | 1 | 0 | 0 | 24 | 4 | 3 | 0 | 2 | 1 |
| DE | 9 | 0 | 0 | 44 | 2 | 1 | 1 | 0 | 9 |
| FL | 17 | 2 | 7 | 192 | 48 | 17 | 28 | 31 | 114 |
| GA | 14 | 0 | 1 | 21 | 10 | 5 | 1 | 7 | 25 |
| HI | 1 | 0 | 1 | 13 | 5 | 1 | 0 | 2 | 2 |
| IA | 10 | 1 | 0 | 29 | 11 | 2 | 1 | 2 | 28 |
| ID | 7 | 0 | 3 | 24 | 4 | 0 | 0 | 7 | 11 |
| IL | 36 | 1 | 2 | 122 | 29 | 28 | 13 | 39 | 154 |
| IN | 6 | 1 | 2 | 28 | 7 | 8 | 3 | 2 | 94 |
| KS | 9 | 0 | 0 | 19 | 7 | 1 | 8 | 2 | 23 |
| KY | 5 | 0 | 1 | 15 | 1 | 14 | 4 | 0 | 24 |
| LA | 2 | 3 | 1 | 11 | 7 | 7 | 12 | 4 | 61 |
| MA | 16 | 18 | 1 | 258 | 37 | 21 | 14 | 37 | 89 |
| MD | 12 | 2 | 3 | 142 | 21 | 14 | 4 | 25 | 26 |
| ME | 1 | 0 | 0 | 10 | 6 | 3 | 3 | 0 | 7 |
| MI | 29 | 1 | 2 | 196 | 36 | 25 | 12 | 14 | 185 |
| MN | 16 | 3 | 2 | 84 | 20 | 13 | 4 | 3 | 57 |
| MO | 15 | 2 | 2 | 57 | 9 | 5 | 2 | 4 | 45 |
| MS | 1 | 2 | 2 | 5 | 0 | 0 | 0 | 3 | 5 |
| MT | 8 | 0 | 0 | 11 | 10 | 1 | 0 | 1 | 1 |
| NC | 2 | 0 | 0 | 54 | 6 | 5 | 1 | 40 | 39 |
| ND | 1 | 0 | 0 | 4 | 1 | 1 | 0 | 0 | 6 |
| NE | 5 | 0 | 0 | 11 | 11 | 0 | 1 | 0 | 7 |
| NH | 4 | 1 | 0 | 32 | 6 | 0 | 1 | 2 | 16 |
| NJ | 13 | 8 | 1 | 234 | 27 | 30 | 17 | 37 | 141 |
| NM | 3 | 1 | 0 | 71 | 17 | 15 | 11 | 17 | 20 |
| NV | 0 | 2 | 1 | 22 | 11 | 4 | 3 | 2 | 8 |
| NY | 39 | 20 | 12 | 337 | 90 | 62 | 20 | 70 | 185 |
| OH | 11 | 1 | 3 | 223 | 17 | 42 | 16 | 33 | 170 |
| OK | 10 | 4 | 1 | 32 | 15 | 8 | 14 | 5 | 111 |
| OR | 3 | 1 | 4 | 27 | 14 | 1 | 1 | 1 | 11 |
| PA | 23 | 5 | 9 | 193 | 31 | 101 | 35 | 329 | 133 |
| PR | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 0 |
| RI | 0 | 1 | 0 | 8 | 3 | 0 | 3 | 0 | 10 |
| SC | 2 | 1 | 0 | 27 | 11 | 4 | 1 | 19 | 14 |
| SD | 0 | 0 | 0 | 7 | 1 | 1 | 0 | 0 | 2 |
| TN | 12 | 0 | 7 | 30 | 5 | 13 | 3 | 37 | 33 |
| TX | 38 | 28 | 4 | 248 | 72 | 49 | 81 | 36 | 531 |
| UT | 8 | 1 | 0 | 32 | 6 | 9 | 4 | 8 | 32 |
| VA | 13 | 0 | 2 | 99 | 31 | 18 | 2 | 42 | 35 |
| VI | 0 | 0 | 0 | 4 | 0 | 1 | 0 | 0 | 0 |
| VT | 1 | 0 | 0 | 6 | 15 | 0 | 2 | 0 | 5 |
| WA | 22 | 1 | 4 | 95 | 34 | 7 | 19 | 28 | 50 |
| WI | 9 | 0 | 1 | 47 | 11 | 5 | 6 | 4 | 91 |
| WV | 6 | 0 | 0 | 8 | 5 | 33 | 4 | 4 | 10 |
| WY | 0 | 0 | 0 | 7 | 7 | 7 | 2 | 0 | 6 |

We observe that bulk energy patenting activity is increasingly coming from non-U.S. sources; the U.S. share of issued clean energy patents has been declining by about 1% per year since 1975 (Figure 4). The recent observed increase in percentage of patents granted to U.S. institutions is likely a mathematical artifact created by comparing the largely declining granted patent trends (Figure 1, Figure 3, and Figure 4). Different countries are responsible for the majority of international patenting, but in general we observe significant activity from Germany, Japan, Canada, and France. Interestingly, three major sources of clean energy activity—India, China, and Brazil [60, 61]—are not even in the top 10 most active patenting countries, although this may be an artifact of their recent entry into the field and USPTO granting delays. In keeping with general features of foreign patents, foreign energy patents have a lower citation count than their U.S. counterparts, presumably because foreign firms build upon their domestic knowledge base and do not spend as much time and effort tracking the U.S. intellectual property landscape [62]. This makes evaluating their impact using patent citations challenging. Future work should include evaluation of a disambiguated global intellectual property database as well as publications to truly measure the innovative capacity of the global economy.

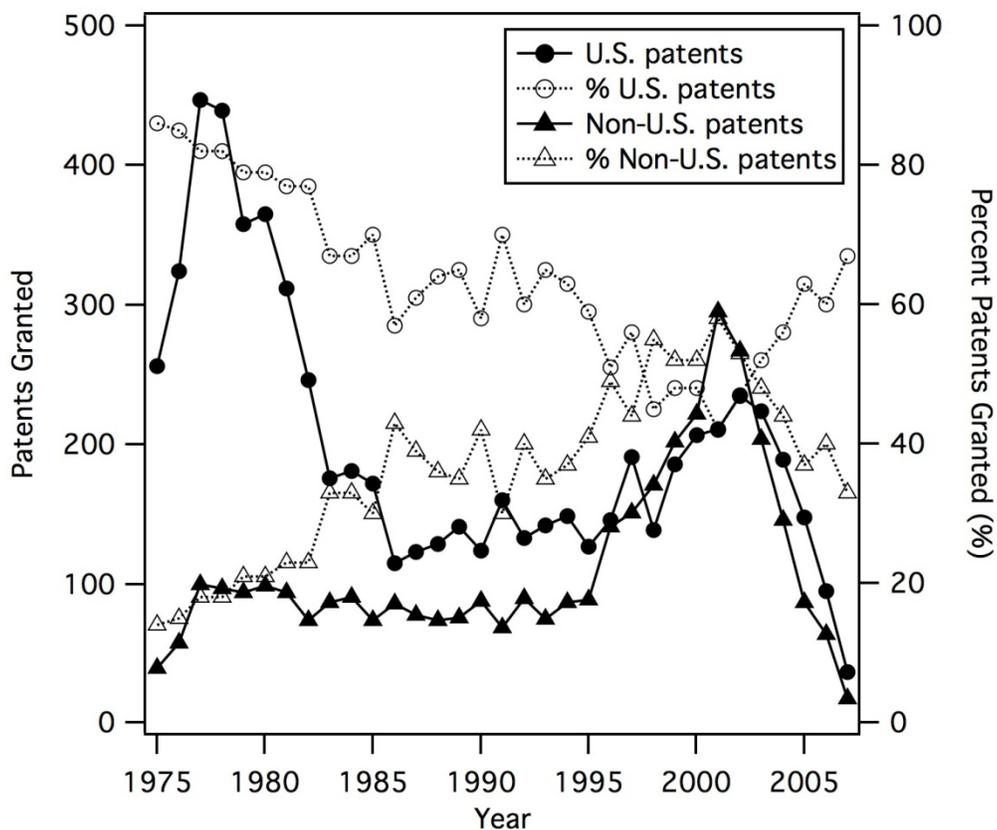


Figure 4. A comparison of renewable energy patents generated by U.S. and non-U.S. inventors

3.1.2 Institutional Patenting Trends

To illustrate the institutional bases of this patenting effort over time, we categorize institutional sources of invention as private firms, government agencies, universities, and independent inventors (the latter are patents without an assignee and are typically owned by the listed inventors) and characterize whence the innovations came (Table 2).

Table 2. Market Sources of Patents

| Org. Type | Energy Technology | | | | | | | | |
|-------------|-------------------|------------|-------|-------|------|-------------|-----|---------|------|
| | Renewable | | | | | Traditional | | | |
| | Biofuels | Geothermal | Hydro | Solar | Wind | Coal | Gas | Nuclear | Oil |
| Private | 370 | 145 | 30 | 2458 | 408 | 489 | 278 | 942 | 2391 |
| Independent | 124 | 71 | 61 | 1974 | 579 | 160 | 130 | 148 | 783 |
| Government | 39 | 7 | 9 | 316 | 57 | 62 | 19 | 95 | 76 |
| Academia | 27 | 3 | 1 | 113 | 16 | 26 | 14 | 53 | 29 |
| Other | 1 | 0 | 0 | 7 | 2 | 1 | 0 | 5 | 7 |

As an example of why this analysis is useful, one might think that wind power is the province of mechanically inclined tinkerers in their garage, while biofuels innovations come from large corporations because of the technical complexity and experimental requirements for discovery. This theory appears to hold true with independent inventors responsible for 54.5% of patents in wind and only 22.1% in biofuels. Hence, independent inventors appear to be a substantial source of technology in wind power and firms in biofuels. A modest proportion of solar patents (33%) also appear to come from independent inventors, which may be counter-intuitive considering the advanced materials chemistry required with most photovoltaic technologies. However, the time series of bulk patenting activity also shows high levels of independent invention during periods when entrepreneurial semiconductor activity was on the rise (1985–1995), and our definition of solar also includes solar thermal technologies that do not require advanced laboratories to fabricate. The technology with the lowest proportion of technologies coming from independent inventors is nuclear (11.9%), which may be due to (1) the inherent dangers of the technology, (2) the lack of market for new innovations because of diminished new construction since the 1980s [63], or (3) the dependence on a limited pool of large corporations for essential equipment manufacturing [64].

Before the early 1980s in solar, independent inventors and corporations were equally dominant patent producers. Since then, however, corporations have produced the majority of patents, with independent inventors contributing only about 25% of patents. From 1995–2008, corporations have extended their dominance, responsible for up to three quarters of the granted solar patents. This transition may be due to the transition of technology dominance from solar thermal to photovoltaics, which are much more experimentally complicated technologies and more the provenance of corporations than of independent inventors.

Wind energy demonstrates a somewhat similar pattern. Independent inventors contributed the most patents in the late 1970s and early 1980s, after which large firms supplanted them. Since then, corporations and independent inventors have contributed about equally to the patent pool. In terms of gross number of patents, universities and national laboratories never figure prominently in prolifically producing wind energy

patents, although we will demonstrate later that universities may, on average, produce more patents with technical breakthroughs.

Patenting in biofuels has been steadily increasing over the time period, and this increase appears to have come mainly from large firms, which maintain a dominant role in patenting activity. This raises the question of whether independent inventors are inventing poor quality technology or whether other issues, such as commercialization challenges or a lack of social networks that diffuse knowledge, are the cause.

3.1.3 Funding Sources

We do not observe broad variability over time in the role of government funding in producing inventive activity. Broadly speaking, in solar, both DOE and the National Aeronautics and Space Administration (NASA) historically account for about one-third of the patents with government rights associated with them. A similar trend is observed in wind and biofuels. The Department of Defense (DOD) and National Science Foundation (NSF) both play a small role in funding patents (< %5 each). It would be interesting to attempt to correlate government technology funding (e.g., Anadon et al. [11]) with patent output. With sufficient data resolution, bibliometric scholars could make some significant evaluations of the effectiveness of funding agencies on stimulating innovation and the efficiency of research and development organizations in converting public funding into public good [65, 66].

3.1.4 Social Features

We also performed preliminary analysis of social networks of patenting. Unfortunately, our keyword-driven search approach creates artificial boundaries around social networks of invention, and for this reason, we do not calculate a unique social influence variable for each inventor since their true social influence should be derived from their entire inventive career, which is typically much broader than this dataset would indicate (Figure 5). In this figure, an NREL innovator is identified (circled), and their energy patents identified in our database are mapped (A). However, that analytical method may underestimate the actual network of collaborators that the inventor has previously worked with (B) and the total network of their inventive sphere (C). The inventor's inventive career likely has spanned other disciplines besides energy, and identifying formative collaborations that lead to energy innovations is essential. Each node is an inventor connected by lines to their collaborators (distance is a network plotting artifact and does not indicate closeness), colors indicate different patent assignees, and square nodes and red lines indicate energy patents versus black lines and circles for non-energy patents.

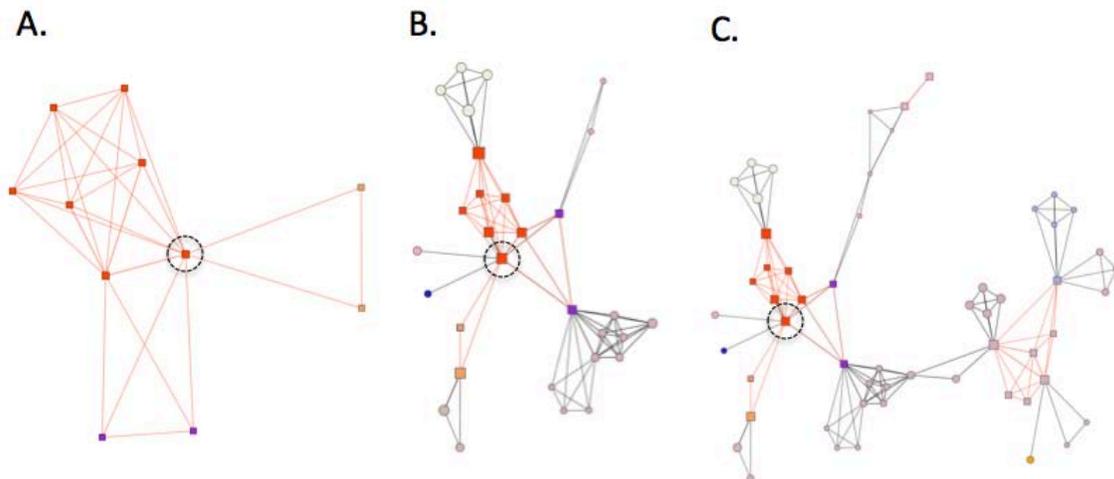


Figure 5. Example collaboration networks for a selected biofuel inventor (dashed circle)

However, we are able to make some preliminary observations about the social features of clean energy patenting. For example, consistent with Wuchty et al. [67], we observe a changing collaborative structure of energy innovation over the time period of study with the typical number of inventors per patent increasing (Figure 6). While the linear trends are not significantly correlated, in the aggregate, we observe team size growing at 2%–3% per year since 1975 for all renewable energy and wind patents.

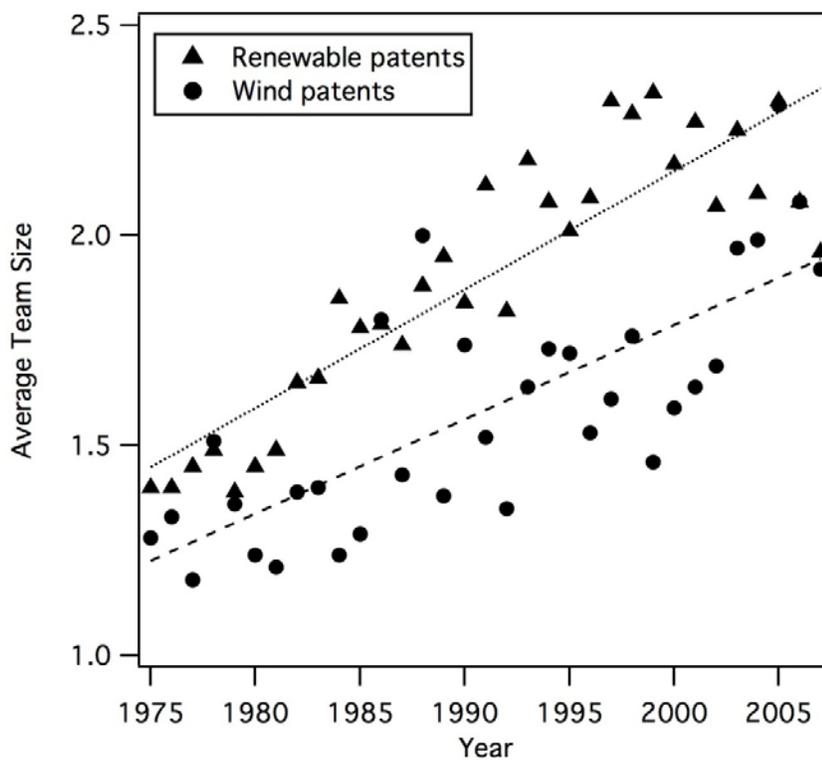


Figure 6. Average collaboration and inventing team size is increasing in energy

Figure 7 shows network dynamics and representations of collaborative networks for all of the energy technologies. For each technology area, we show the complete inventive network (excluding non-collaborating inventors), the largest component of each network, and the network of the most highly cited inventor (node and line coloring is described in Figure 5). As noted above, accurate quantitative descriptions of inventive social networks demand a level of technology-specific boundary robustness that is beyond the scope of this preliminary examination. Additionally, inventor-level social influence can only be accurately characterized by analyzing the social network signature across the full USPTO dataset. But even though our datasets exhibit artificial social network boundaries, preliminary results suggest significant collaborative network differences between the technology fields at both the macro and micro levels. At the macro level, biofuels patenting activity is more islanded with few inter-organizational linkages connecting network components. Wind patenting activity sees moderate levels of inter-organizational linkage, and solar sees by far the highest level of the three technology areas.

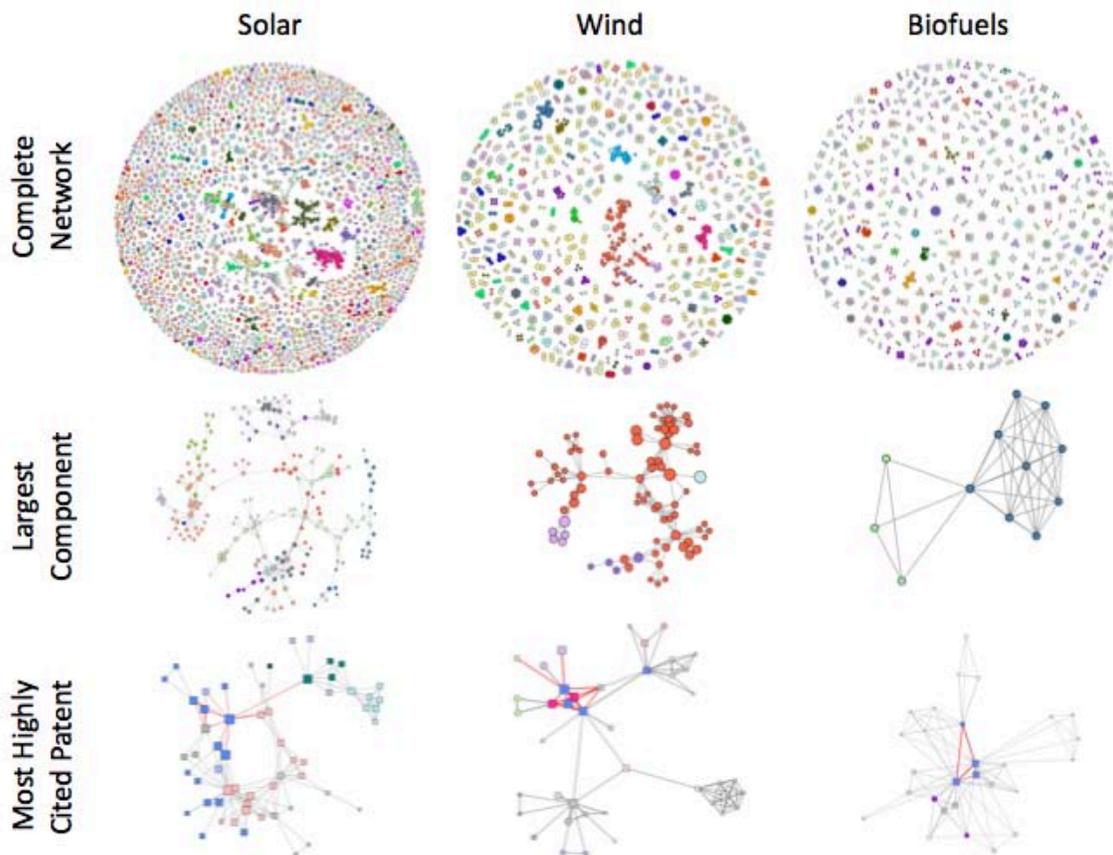


Figure 7. Summary network visualizations for each technology area

As a preliminary test of whether these macro levels of isolation are a genuine network feature or an artifact of the bounded dataset, we isolated the immediate inventive

networks of the highest cited inventors in each technology-bounded dataset and compared these networks to the same inventors' "second-degree" networks extracted from the entire USPTO dataset (Figure 7). In this way, we can gain a first approximation of the full extent of social influence of top inventors (Figure 5). Comparing the networks of the largest components and most highly cited inventors across technology areas suggest differences in career connectivity of inventors in the different technology areas (Figure 7), hinting at fruitful avenues for future research. Furthermore, we did a preliminary comparison of the energy inventive space with that of software and pharmaceuticals and observe massive differences in the collaboration networks. Here we focus on energy patents so this thread of inquiry is not continued, but there are lessons to be learned from these highly collaborative and innovative industries, which will be the focus of future analysis.

Certainly, the historical sources of innovation, the industrial roots, and the resource intensity of these technology areas affect these collaboration norms. Validating these disparate network connectivity characteristics and comparing them to other industries will require more extensive patent datasets, but we suspect that there are indeed substantive differences between the technology fields and important lessons to be learned.

3.2 Sources of Clean Energy Technical and Commercial Breakthroughs

The count-based methods used in this first analytical effort do not account for patent quality, which misses most of the impact of patents, as the distribution of commercial value is extremely skewed [68]. It is very difficult to predictively quantify the impact of these patents, and this quandary is almost universal for bibliometric scholars of research and innovation. While recent databases of patents and papers [42, 69] have greatly facilitated large-scale analyses, they stop short of providing a comprehensive picture of the later stages of application and commercialization.

Breakthrough innovations are fundamentally new technologies or services that change the field of practice. This dramatic shift is relative in comparison to incremental innovations that improve existing technologies and relatively useless innovations that do not significantly advance the technical or commercial practice. Breakthrough energy innovations have been deemed necessary to meaningfully compete with incumbent technologies [70], overcome the significant market and political barriers to system transformation [71], and play a meaningful role in mitigating climate disruption [72]. At the same time, these innovations are likely to be uniquely difficult to achieve from a technical perspective [73]. Given the disproportionate impact of breakthroughs on technical advance and societal change [33, 34, 42], we now focus on these most technically and commercially important patents as measured by citations and Web presence, respectively. As previously mentioned, citations have a long history in the bibliometric literature, and we will use them to identify and quantify technical breakthroughs. However, patent citations are a limited measure of technical importance and do not necessarily infer commercial relevance or likelihood of deployment.

Supplementing the use of patent citations, we propose to assess commercial value through automated searches of the Web and content analysis of the resulting URLs. We

demonstrate the proof-of-concept herein using energy patents and propose it as a candidate tool for general use for the entire patent record. This variable will add an entirely new measure that is systematically calculable across all patents to the toolkits of innovation researchers. This work brings innovation researchers—in clean energy as well as other industries—closer to a measure of deployment and commercial impact.

3.2.1 Web Hit Analysis

We propose that websites and their content can contain valuable data on the value of patents that is supplemental to citation data. The presence of a description of a patent on a non-aggregating website requires qualitative human intervention (i.e., someone needs to create original content that elevates the patent's unique Web presence). This investment of time and author's perceived value of the patent [74] is captured and codified in the Web page. We can then count those Web pages and quantify a perception of unique value.

For example, U.S. Patent 6,852,920, “Nano-architected/assembled solar electricity cell” elicits 209 Web hits. While most of these URLs are patent aggregator results, others provide an indication of commercial interest, such as the following⁴ portion of a Wikipedia entry on the firm that owns the patent (*italics added*):

The company uses copper indium gallium diselenide—which achieves up to 19.9% efficiency in laboratory samples^[13]—to build their thin film solar cells. The company's technology gained early industry recognition with the presentation of a Small Times Magazine award at a leading nanotech business event in 2005.^[14] Nanosolar's solar cells have been verified by NREL to be as efficient as 14.6% in 2006^[15] and 15.3% in 2009.^[6] *Technical details of Nanosolar's new manufacturing techniques have been disclosed in patent applications.*^[17] Some information about their process has become available in a Scientific American article (in German).^[15]

Citation number 17 (other citations are Wikipedia references and omitted for brevity) cites U.S. Patent 6,852,920 and thus provides the link to evidence that the patent is commercially valuable and being actively developed. This Web page has been “peer reviewed” by the open-source community [75]: the page was edited by 147 unique users 286 times from 2006 to 2010 with the most active contributor providing 12 edits to the page.

Conversely, aggregator Web sites generally only replicate the USPTO database information, e.g. patent title, abstract, inventors, and application and grant dates, and provide no new signal on the value of a patent. Figure 8 illustrates systematic relationships between the number of future prior art citations and the number of Web hits for solar patents in 1995 (A) and in 2005 (B). As can be seen, the relationship between citations and Web hits is positive for both aggregator and signal URLs. In the 1995 patents (8A) there exists a significant correlation between cites and aggregator URLs

⁴ Wikipedia. “Nanosolar.” <http://en.wikipedia.org/wiki/Nanosolar>. Accessed December 15, 2010.

($\omega_{agg} = 4.3 \pm 0.1(\gamma) + 41.1 \pm 2.5$, $p < 0.1$), which results from the prior art fields in the aggregator websites and provides little new information on the value of the patent even though the signal is amplified.

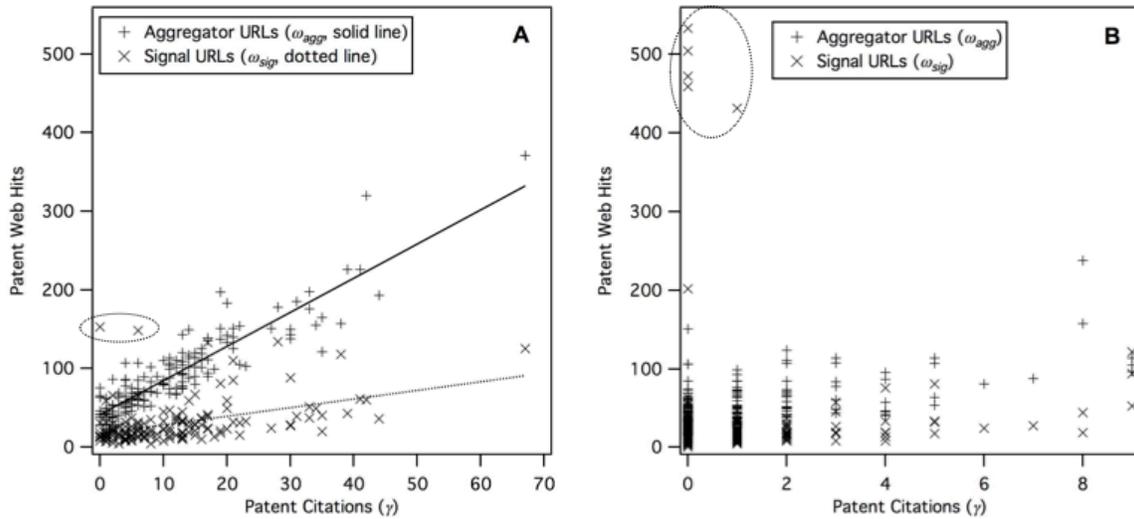


Figure 8. Web hits versus future prior art citations from patents in 1995 (A) and 2005 (B) solar energy patents

However, signal websites ($\omega_{sig} = \omega - \omega_{agg}$) can provide unique information that diverges from the predominantly linear correlation observed in aggregator Web hits because their Web presence is the result of human intervention and is a measure of market importance or commercial interest. Indeed, the signal Web hits contain significantly more variability and do not correlate ($\omega_{sig} = 1.1 \pm 0.2(\gamma) + 17.4 \pm 2.8$, $p > 0.5$). Interesting outliers can be identified when the patents have a significantly higher ratio of signal Web hits to citations, such as those patents that are circled by the dotted lines in Figure 8. These data points highlight Web information, which may be unique compared to prior art information. Qualitative inspection of the highlighted patents shows that the Web hits result from an active description similar to the aforementioned Wikipedia example. Only one of seven highlighted patents has a high signal count that comes from a true error wherein the search terms occur in websites; these websites that use the same terms as designations for an unrelated company. These data indicate that while the measure is nuanced and requires some manual validation, the signal contained in these Web hits is different from the citation data and worth further investigation.

To extend this initial investigation, we also compared two populations of patents, those that were licensed ($n = 81$) and those that were not ($n = 304$). An assumption is that licensed patents are of more commercial or market importance than the unlicensed patents [76]. We expect that they may have a different signature of Web hits as a measure of that commercial impact. Each patent received a mixture of aggregator (B) and signal (A) Web hits (Figure 9), which appeared to have different signatures for licensed versus unlicensed patents.

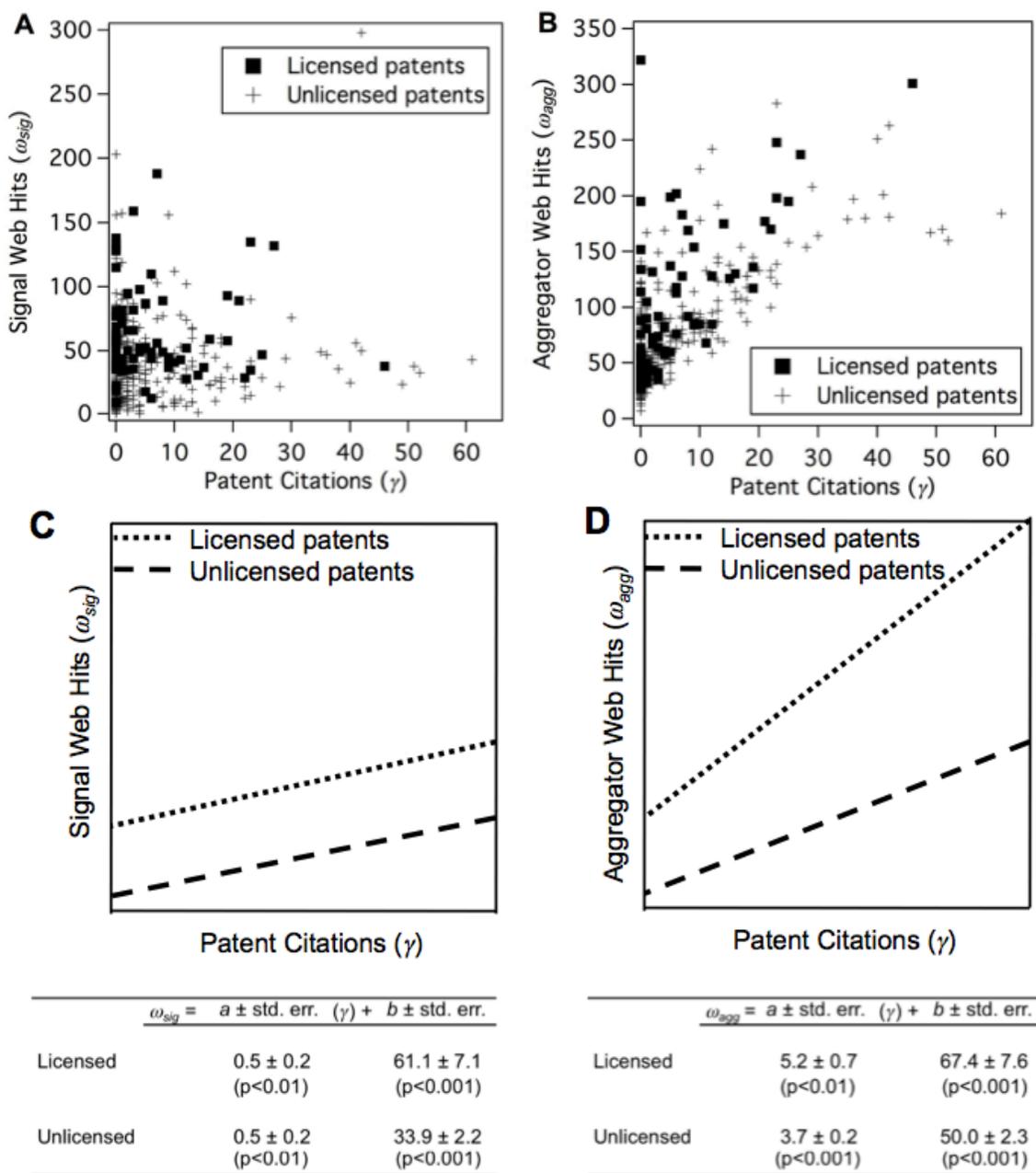


Figure 9. Signal (A, C) and aggregator (B, D) Web hits versus future prior art citations from patents for licensed and unlicensed National Renewable Energy Laboratory patents

To determine if there is a significant difference between the citation and Web hit signatures of the licensed versus unlicensed patents, the patents were analyzed using a simple regression that incorporated a dummy variable to indicate if each patent was licensed and determined the effect of being licensed on the correlation between citations (γ) and Web hits (ω):

$$\omega_{agg,sig} = k_1(\gamma) + k_2(\lambda) + k_3(\gamma\lambda) + b$$

where λ defines if the patent is licensed (licensed = 1, unlicensed = 0), b is the y-axis intercept, and the k 's are constants. We observe significant differences in the signals of these patent populations—calculated results and stylized representations of the data for signal and aggregator Web hits results are presented (Figure 9C and 9D, respectively). While both scale linearly with patent citations, licensed patents receive more signal Web hits per citation than unlicensed patents. Licensed patents also appear to accumulate aggregator Web hits to a greater degree per citation than unlicensed patents. This acceleration of accumulating Web hits in licensed patents may allow this measure to appear more quickly than citations because of the delay in patent issuance [42], thus representing a harbinger of technical importance.

Scholars use patent citation analysis to retrospectively assess the patent record to understand modes of collaboration, analogous to the work done in this study. However, this Web presence indicator may allow innovators, scholars, and research managers to obtain more dynamic and real-time intelligence regarding the commercial value of patents. For example, two unique patents—one commercially used and the other not—are indiscernible in terms of vintage and number of citations but can exhibit significantly different Web presence signatures. Since the mechanics of Web presence accumulation are materially different from those underlying citations, this novel measure provides new information on the value of a patent. Populations of licensed and unlicensed patents are known to be different given the *ex post* information about whether they have generated commercial interest, and we demonstrate that the populations exhibit different signatures in the aggregator and signal Web hits that are statistically different. The differences in the signal Web hits are the expected differences in the commercial impact of the patents while the aggregator Web hits are likely an amplified measure of the technical impact of the patents, typically measured by citations. With further refinement, aggregator Web hits may be able to mirror analysis that currently uses citations to predict technical impact, while adding signal Web hits as a measure of commercial importance. Combination of these metrics may allow for easier and faster assessment of patent impact. While this measure is noisy and requires refinement, these preliminary data indicate that Web presence analysis may be able to help determine or predict commercial value of patents as distinguished from technical merit.

In an attempt to extend these observations that signal Web hits are a measure of commercial importance, we performed the same Poisson regressions using signal hits that we performed on the patent citations to test differences between the factors producing commercially relevant, rather than technically relevant, innovations.

3.3 Breakthrough Sources

Technology importance (citation based) and commercial importance (Web presence based) findings are summarized by technology in the following sections. Effect sizes are expressed in terms of absolute citation or Web hit count differences, and we report both observed and simulated variance of like patents. Observed mean values for the patent group of interest (b) are compared against all available control groups (a) for the contrast variable of interest. We report all significant findings (regressed or simulated $p < 0.1$); greater coefficient size and agreement between the two statistical methods indicates increased certainty. For each reported effect, there is an unreported significant finding in

the opposite direction ($a - b$). For example, a simulated difference of +1.5 for Texas biofuels patents over New York patents (Table 5, line 2) denotes that we expect that patents from Texas receive 1.5 more citations than patents from New York (i.e., patents from Texas are on average 30% more highly cited). Inversely, New York biofuels patents are 24% less highly cited and therefore less technologically important than patents from Texas. Data herein is preliminary and represents a limited set of identified energy patents. However, it does provide interesting evidence of differences between technology areas and paths for future exploration in clean energy patenting and innovation. In the following sections, we provide descriptions of the sources of technical and commercial breakthrough patents in the solar, wind, and biofuels technology areas.

3.3.1 Solar

For technical importance, unassigned solar patents from independent inventors are significantly less cited than patents owned by universities, national labs, and corporations [29] (Table 3 for all solar data). Drawing on the previous discussion (§3.1.2.), while most solar patents came from corporations compared to independent inventors, it appears that those patents from corporations were more technically important. There likely may be some data artifacts here due to the technological differences between photovoltaics and solar thermal innovations, but the differences between sources of total patents compared to breakthrough patents is still relevant. Solar patents owned by corporations and universities are the most highly cited patents, with an insignificant citation increase to university patents. We do not observe any strong relationship to funding sources. Increasing team size correlates with increasing technical breakthroughs. In terms of geographic effects, patents from California and Massachusetts are both highly cited. The effect in California is likely due to the predominance of the semiconductor industry and resultant solar-applicable technologies that are heavily concentrated there, [77] while the effect in Massachusetts is more likely driven by universities and a few highly innovative solar firms (e.g., Konarka and Evergreen Solar). Colorado patents are less highly cited relative to patents from other states. This observation may be due to the higher percentage of solar patents in Colorado originating from independent inventors than in the whole solar dataset (46% versus 33%, respectively). Singh and Fleming [29] have previously demonstrated that independent inventors often face challenges in bringing their technologies to fruition. However, the patents from these independent inventors enjoy an average citation count higher than the patents from the national laboratories (6.7 versus 6.1, respectively), indicating that perhaps these patents are more technically relevant.

Table 3. Regressed and Simulated Determination of Sources for Solar Technical and Commercial Breakthrough Innovation

| Outcome Measure | Contrast Variable ¹ | Contrast From (a) | Contrast To (b) | Mean Outcome | | Observed Mean Difference (b-a) | Simulated Difference (b-a) | Regressed p Value | Simulated p Value |
|-----------------|--------------------------------|-------------------|-----------------|--------------|--------|--------------------------------|----------------------------|-------------------|-------------------|
| | | | | From (a) | To (b) | | | | |
| Citations | State | CO | CA | 6.153 | 7.881 | 1.729 | 1.409 | 0.025 * | 0.000 *** |
| Citations | State | CO | MA | 6.153 | 9.976 | 3.824 | 1.714 | 0.065 * | 0.002 ** |
| Citations | State | CO | NY | 6.153 | 8.239 | 2.087 | 1.263 | 0.002 ** | 0.007 ** |
| Citations | State | CO | TX | 6.153 | 8.450 | 2.298 | 1.333 | 0.024 * | 0.007 ** |
| Citations | State | CO | Misc | 6.153 | 8.352 | 2.199 | 0.950 | 0.002 ** | 0.009 ** |
| Citations | State | NY | CA | 8.239 | 7.881 | -0.358 | 0.146 | 0.047 * | 0.359 |
| Citations | State | Misc | CA | 8.352 | 7.881 | -0.470 | 0.458 | 0.014 * | 0.017 * |
| Citations | State | Misc | MA | 8.352 | 9.976 | 1.624 | 0.741 | 0.258 | 0.056 * |
| Citations | Funding | DOE | NASA | 7.902 | 9.432 | 1.530 | 0.130 | 0.076 * | 0.423 |
| Citations | Funding | NASA | O-Govt | 9.432 | 8.805 | -0.627 | 0.202 | 0.006 ** | 0.380 |
| Citations | Funding | Other | O-Govt | 8.150 | 8.805 | 0.655 | 0.657 | 0.010 ** | 0.096 * |
| Citations | Assignee | Prvt | Univ | 9.088 | 9.142 | 0.054 | 0.404 | 0.014 * | 0.247 |
| Citations | Assignee | Govt | Prvt | 9.122 | 9.088 | -0.034 | 1.035 | 0.626 | 0.003 ** |
| Citations | Assignee | Govt | Univ | 9.122 | 9.142 | 0.020 | 1.466 | 0.041 * | 0.017 * |
| Citations | Assignee | None | Prvt | 6.939 | 9.088 | 2.148 | 1.672 | 0.000 *** | 0.000 *** |
| Citations | Assignee | None | Univ | 6.939 | 9.142 | 2.202 | 2.114 | 0.175 | 0.000 *** |
| Citations | Assignee | None | Govt | 6.939 | 9.122 | 2.182 | 0.652 | 0.161 | 0.048 * |
| Citations | Inventor | One Inv | Two Inv | 7.695 | 8.988 | 1.293 | 0.887 | 0.372 | 0.000 *** |
| Citations | Inventor | One Inv | Three+ | 7.695 | 9.260 | 1.565 | 1.370 | 0.903 | 0.000 *** |
| Citations | Inventor | Two Inv | Three+ | 8.988 | 9.260 | 0.272 | 0.479 | 0.493 | 0.061 * |
| Web Hits | State | CA | CO | 23.351 | 27.350 | 3.999 | 3.209 | 0.562 | 0.094 * |
| Web Hits | State | MA | CO | 23.412 | 27.350 | 3.938 | 4.405 | 0.697 | 0.072 * |
| Web Hits | State | Misc | CO | 23.824 | 27.350 | 3.525 | 4.257 | 0.438 | 0.033 * |
| Web Hits | State | Misc | TX | 23.824 | 25.816 | 1.992 | 2.947 | 0.171 | 0.088 * |
| Web Hits | Funding | Other | NASA | 23.617 | 27.719 | 4.102 | 3.562 | 0.793 | 0.063 * |
| Web Hits | Assignee | Prvt | Univ | 24.051 | 27.965 | 3.914 | 4.556 | 0.389 | 0.022 * |
| Web Hits | Assignee | Govt | Univ | 25.240 | 27.965 | 2.724 | 4.624 | 0.488 | 0.043 * |
| Web Hits | Assignee | None | Univ | 23.625 | 27.965 | 4.339 | 5.013 | 0.297 | 0.013 * |
| Web Hits | Inventor | One Inv | Three+ | 23.944 | 25.716 | 1.773 | 2.328 | 0.558 | 0.016 * |
| Web Hits | Inventor | Two Inv | One Inv | 23.159 | 23.944 | 0.784 | 0.591 | 0.096 * | 0.270 |
| Web Hits | Inventor | Two Inv | Three+ | 23.159 | 25.716 | 2.557 | 2.936 | 0.040 * | 0.005 ** |

¹State designations: According to convention except Misc, which is all states excluding CA, MA, NY, TX, and CO. Funding designations: DOE, Department of Energy; NASA, National Aeronautic and Space Administration; DOD, Department of Defense; O-Govt, Other government agencies, (e.g., National Science Foundation); Other, All patents that did not have government rights identified in the patent. Assignee designations: Univ, Academia; Prvt, Corporations; Govt, Government and National Laboratories; None, Independent. Inventor designation: One Inv, a single named inventor; Two Inv, Two named inventors; Three+, Three or more named inventors.

In solar, we generally observe weak signal significance for all findings of commercial importance. Colorado creates more commercially relevant solar patents than all other states. We observe a higher ratio of Web hits to citations from government patents than these independent patents (8.9 versus 6.0, respectively), which may be the result of truly higher commercial importance. However, the National Renewable Energy Laboratory advertises its available intellectual property on a website, the Energy Innovation Portal,⁵ which may skew the results. Determining the sources of these differences in signal patterns warrants further investigation and again is essential for validating this methodology.

⁵ DOE. "Energy Information Portal." <http://techportal.eere.energy.gov/>. Accessed February 16, 2011.

Texas appears to produce more commercially relevant solar patents than most other states. NASA appears to fund commercially important research, while no other government agencies do in this dataset. Universities dominantly produce more commercially impactful solar innovations. Again, increasing team size appears to produce more commercially impactful patents.

These preliminary data hint at a range of potential prescriptive actions that could be taken to accelerate the development of breakthrough innovations in solar. First, public and private research could focus on funding teams of researchers rather than individuals. Previous research tells us that independent inventors typically produce more failed inventions but also can create more disruptive breakthroughs because independent innovators do not have collaborators inducing “groupthink” [78, 79]. Despite the promise of radical breakthroughs, independent inventors on average produce lower value innovations and may hinder efforts to optimize research dollars toward increasing the likelihood of major impact. Alternately, perhaps, independent innovators should continue to be allowed to innovate independently but then be provided with better mechanisms to support technology commercialization through collaboration with large organizations; e.g., the NREL Commercialization Assistance Program (NCAP).⁶ Research increasingly shows that breakthrough innovations can be accelerated in environmental contexts that maximize cross-pollination and the free flow of ideas [79]. Our results provide quantitative insight on the effectiveness of particular geographic clusters and indicate which of these may be exploited. The data suggest that California may excel at producing important but incremental innovations, while Colorado may produce more commercially valuable innovations. While the institutional and social sources of these differences merit further investigation, innovation clusters are increasingly considered to be path-dependent: accumulated human, physical, and financial capital in a particular technology sector predisposes the cluster to future innovation in that sector [80, 81]. Building on this observation, universities with historically important solar innovations should be looked to as breeding grounds for new commercially impactful breakthroughs, while corporations may be better at creating technical breakthroughs that incrementally enrich the technological landscape of the whole industry. Finally, on a more qualitative bent requiring further research, it is apparent that NASA has effectively sponsored high impact research from both a technical and commercial perspective. Qualitative assessment of their funding priorities, process, and management may provide additional insights on how to improve other public departments’ efforts.

3.3.2 Wind

DOE- and NASA-funded wind patents are cited more strongly than other government-funded wind patents (Table 4 for all wind data). Larger teams are associated with more highly cited patents. Texas wind patents are less highly cited than those from other states. It appears that this is due to a high percentage of independent inventors (55% versus 49% average) that produce uncharacteristically low value patents (average independent Texas citations = 4.3 versus average independent citations for all wind patents = 5.9). Wind

⁶ NREL. “Commercialization Assistance Program,” <http://www.nrel.gov/technologytransfer/ncap.html>. Accessed March 7, 2011.

patents from corporations appear to be the most highly cited. As previously noted, the number of collaborators per patent has been steadily increasing in wind and renewable energy [29, 67] (Figure 6), and this trend also appears to positively affect the likelihood of breakthroughs in wind. While we cannot currently extrapolate how multidisciplinary the teams are [37], this finding indicates that work to find breakthrough technical innovations should be purposefully collaborative.

For commercial breakthroughs, Colorado produces significantly more commercially relevant wind patents. This effect may be due in part to the presence of the National Renewable Energy Laboratory, which has produced 18% of the wind patents in Colorado. However, the more commercially impactful wind patents come from independent inventors: independent patents receive more than twice as many signal Web hits than the average for all wind patents. The wind sector also appears to be unique in that independent inventors excel in both technical and commercial impact.

Table 4. Regressed and Simulated Determination of Sources for Wind Technical and Commercial Breakthrough Innovation

| Outcome Measure | Contrast Variable ¹ | Contrast From (a) | Contrast To (b) | Mean | Mean | Observed Mean Difference (b-a) | Simulated Difference (b-a) | Regressed p Value | Simulated p Value |
|-----------------|--------------------------------|-------------------|-----------------|------------------|----------------|--------------------------------|----------------------------|-------------------|-------------------|
| | | | | Outcome From (a) | Outcome To (b) | | | | |
| Citations | State | TX | CA | 4.985 | 6.990 | 2.005 | 2.088 | 0.066 * | 0.017 * |
| Citations | State | TX | CO | 4.985 | 8.851 | 3.866 | 2.728 | 0.081 * | 0.015 * |
| Citations | State | TX | NY | 4.985 | 7.203 | 2.218 | 2.631 | 0.004 ** | 0.012 * |
| Citations | State | TX | Misc | 4.985 | 6.865 | 1.880 | 1.437 | 0.004 ** | 0.046 * |
| Citations | State | Misc | NY | 6.865 | 7.203 | 0.338 | 1.176 | 0.279 | 0.094 * |
| Citations | Funding | O-Govt | DOE | 6.793 | 5.634 | -1.159 | 2.540 | 0.252 | 0.095 * |
| Citations | Funding | O-Govt | NASA | 6.793 | 10.056 | 3.262 | 2.631 | 0.629 | 0.091 * |
| Citations | Funding | Other | DOE | 6.868 | 5.634 | -1.234 | 2.749 | 0.024 * | 0.028 * |
| Citations | Funding | Other | NASA | 6.868 | 10.056 | 3.188 | 2.843 | 0.163 | 0.019 * |
| Citations | Assignee | Govt | Prvt | 8.778 | 6.811 | -1.967 | 2.041 | 0.288 | 0.024 * |
| Citations | Assignee | None | Prvt | 6.823 | 6.811 | -0.012 | 0.925 | 0.783 | 0.025 * |
| Citations | Assignee | None | Univ | 6.823 | 7.188 | 0.364 | 1.073 | 0.096 * | 0.262 |
| Citations | Inventor | One Inv | Two Inv | 6.643 | 8.150 | 1.507 | 1.475 | 0.273 | 0.006 ** |
| Citations | Inventor | One Inv | Three+ | 6.643 | 6.612 | -0.031 | 2.741 | 0.208 | 0.000 *** |
| Citations | Inventor | Two Inv | Three+ | 8.150 | 6.612 | -1.538 | 1.237 | 0.770 | 0.069 * |
| Web Hits | State | CA | CO | 20.717 | 38.479 | 17.762 | 14.109 | 0.012 * | 0.001 *** |
| Web Hits | State | MA | CO | 23.818 | 38.479 | 14.661 | 14.700 | 0.053 * | 0.003 ** |
| Web Hits | State | NY | CO | 20.886 | 38.479 | 17.593 | 15.675 | 0.004 ** | 0.000 *** |
| Web Hits | State | TX | CO | 19.227 | 38.479 | 19.251 | 17.692 | 0.006 ** | 0.000 *** |
| Web Hits | State | Misc | CO | 25.213 | 38.479 | 13.266 | 13.628 | 0.021 * | 0.000 *** |

¹Contrast Variable definitions: see Table 3.

3.3.3 Biofuels

We found that government funding had a significant impact on citations as a measure of highly important technical innovations (Table 5 for all biofuels data). DOE-funded research that resulted in biofuels patents was less cited relative to “other” government sources. Similar to DOE, NASA biofuels patents suffered a penalty compared to other government funding sources. Inventor teams did not benefit from greater size in biofuels, which runs counter to recent research [29, 67] and our findings in wind; this anomaly bears investigation. This analysis did not observe any strong geographic effect on where breakthroughs are coming from, except to note that patents from New York appear to have slightly lower citation rates. More than half of the 40 biofuels patents in the group

from New York were invented by two people from Texaco Inc. and were related to fuel additives or mixtures. We cannot rationalize the observed citation penalty from these data but suggest that these particular inventions play a role in weighing down the citation count in this geography. Finally, we observe a weakly significant signal for university patents being more highly cited than inventions from independent inventors; no other significant effects from inventing institutions are observed.

Table 5. Regressed and Simulated Determination of Sources for Biofuel Technical and Commercial Breakthrough Innovation

| Outcome Measure | Contrast Variable ¹ | Contrast From (a) | Contrast To (b) | Mean | Mean | Observed Mean Difference (b-a) | Simulated Difference (b-a) | Regressed p Value | Simulated p Value |
|-----------------|--------------------------------|-------------------|-----------------|------------------|----------------|--------------------------------|----------------------------|-------------------|-------------------|
| | | | | Outcome From (a) | Outcome To (b) | | | | |
| Citations | State | CA | Misc | 6.063 | 7.076 | 1.012 | 0.205 | 0.073 * | 0.345 |
| Citations | State | NY | TX | 7.184 | 9.395 | 2.211 | 1.510 | 0.364 | 0.058 * |
| Citations | State | NY | Misc | 7.184 | 7.076 | -0.108 | 0.984 | 0.137 | 0.071 * |
| Citations | Funding | DOE | O-Govt | 3.360 | 8.813 | 5.453 | 4.036 | 0.018 * | 0.013 * |
| Citations | Funding | DOE | Other | 3.360 | 7.293 | 3.933 | 3.440 | 0.036 * | 0.012 * |
| Citations | Funding | NASA | O-Govt | 3.556 | 8.813 | 5.257 | 3.112 | 0.014 * | 0.040 * |
| Citations | Funding | NASA | Other | 3.556 | 7.293 | 3.737 | 2.524 | 0.018 * | 0.031 * |
| Citations | Assignee | None | Univ | 8.150 | 8.704 | 0.554 | 1.212 | 0.287 | 0.092 * |
| Web Hits | State | MA | CA | 15.900 | 23.905 | 8.005 | 10.234 | 0.051 * | 0.015 * |
| Web Hits | State | MA | CO | 15.900 | 24.615 | 8.715 | 10.720 | 0.316 | 0.021 * |
| Web Hits | State | MA | NY | 15.900 | 34.316 | 18.416 | 17.949 | 0.045 * | 0.015 * |
| Web Hits | State | MA | TX | 15.900 | 27.592 | 11.692 | 14.738 | 0.058 * | 0.028 * |
| Web Hits | State | MA | Misc | 15.900 | 23.360 | 7.460 | 10.704 | 0.015 * | 0.008 ** |
| Web Hits | Assignee | Prvt | Univ | 22.681 | 27.222 | 4.541 | 4.756 | 0.516 | 0.066 * |
| Web Hits | Assignee | Prvt | Govt | 22.681 | 27.421 | 4.740 | 5.149 | 0.126 | 0.065 * |

¹ Contrast Variable definitions: see Table 3.

For commercial breakthroughs, the only geographic effect that we observe is that biofuels patents from Massachusetts are highly discounted in their commercial impact. The single largest assignee of biofuels patents from Massachusetts is the Massachusetts Institute of Technology (MIT) ($n = 4$) with a single inventor responsible for all these patents. MIT is highly lauded as an innovation and entrepreneurship center, and we observe a strong Web hits signal for these patents—in fact, one of these patents is also the most highly cited—so we expect the observed penalty is not from these patents. The remaining patents in the dataset are for a diverse group of corporations and independent inventors.

There are some important differences in the intellectual property practices of different market players that need to be properly accounted for to improve this measure of Web presence. The observation of a weak Web hit signal from corporations may be the result of two related but distinct factors: (1) Universities must “advertise” their inventions to get them licensed and into commercial practice, which could increase their Web hits measure, (2) Corporations may purposefully not advertise their innovations and, if they are a small company without much outside interest in the blogosphere, may avoid detection by the Web community, decreasing their Web hits signal. Determining which of these factors are at play requires further investigation and is essential to the correct use of the Web hits measure but is beyond the scope of this paper. There is a strong but not very significant observation that biofuels patents from universities and national laboratories are more commercially relevant. We do not observe any effect of funding or team size on the commercial importance of biofuels patents.

These empirically derived findings contribute to the dialogue on where and how innovations, and energy innovations in particular, originate and how we can create more of them. In particular, it has been widely accepted that collocated communities of practice—universities, corporations, and venture capital—are helpful and perhaps essential for creating vibrant innovation ecosystems [82, 83]. Indeed, in many industries, such as those related to energy including biotechnology (biofuels) and semiconductors (solar), the areas of Silicon Valley [84, 85] and Greater Boston [82, 86] are indicated as innovation and company hotbeds. However, these data demonstrate that identified biofuels patents from California are mildly discounted when compared to all other patents. Similarly, the patents from Massachusetts, while perhaps technically important, have limited commercial importance. We realize the limitations of these interpretations and hesitate to extend these observations much further without significantly more investigation. However, unique technology development and deployment dynamics exist within the energy sector due to the nature of the technologies and markets. These unique differences require us to reevaluate how public and private research, development, and deployment capital is used.

Biofuels is our smallest dataset, with only 938 patents. Each of the geographical, social, and institutional patent sub-populations is proportionally limited, so finding statistically significant correlations between the patent populations is more difficult than in the solar and wind sets. Furthermore, the work of prescribing reliable funding strategies becomes more tenuous as the sample size decreases, so we will focus our preliminary prescriptive efforts on the richest dataset available—solar patents. However, we expect to be able to add prescriptive detail as patent sample sizes increase through improved search methods.

3.4 Predicting Breakthroughs

Through a retrospective analysis of energy patents, we have demonstrated the ability to ascertain key drivers of breakthrough technical and commercial impact. The introduction of the new measure of screened Web hits as a predictor of commercial value shows significant promise as a tool for bibliometric researchers. The logical next step is to determine if this type of analysis can be used predicatively to anticipate breakthrough or commercially impactful innovations. While we are admittedly a long way from being able to do this with any assurance of accuracy, we define the following equation in an attempt to begin a thread of inquiry wherein scholars and practitioners alike can anticipate the value of a patent. For the licensed and unlicensed NREL patents, we calculate the following ratio for each one:

$$\kappa = (\omega_{sig} - \bar{\omega}_{sig}) + (\omega_{agg} - \bar{\omega}_{agg})$$

where κ is the calculated potential impact of the patent, and the average signal or aggregator Web hit value is subtracted from the unique patent Web hit value. If κ for a patent is positive, that patent is more likely to be impactful or commercially relevant. Indeed, for our two populations of patents, 62% of the licensed patents are correctly tagged positive while only 30% of the unlicensed patents are. Furthermore, comparing the average commercial potential for the licensed ($\bar{\kappa}_{lic}$) versus unlicensed ($\bar{\kappa}_{unlic}$) patents by year shows that the relationship predominantly correctly predicts commercial value

for the licensed patents within a given year (Figure 10) while accounting for truncation in both citations and Web hits. For each year, the calculated variable generally correctly approximates commercial value of the licensed patents higher than the unlicensed patents, i.e. $\bar{\kappa}_{lic} - \bar{\kappa}_{unlic} > 0$. Similar to citations (Figure 1), Web presence appears to suffer a degree of signal truncation, although less severe than citations: patents from the last decade average at least 15–20 signal Web hits. Hence, albeit with significant future refinements, we believe that it is possible to construct a predictive relationship using the Web hits metric. Furthermore, this measure may be a complement to citation analysis and may appear earlier than those records due to the delay in citation evolution.

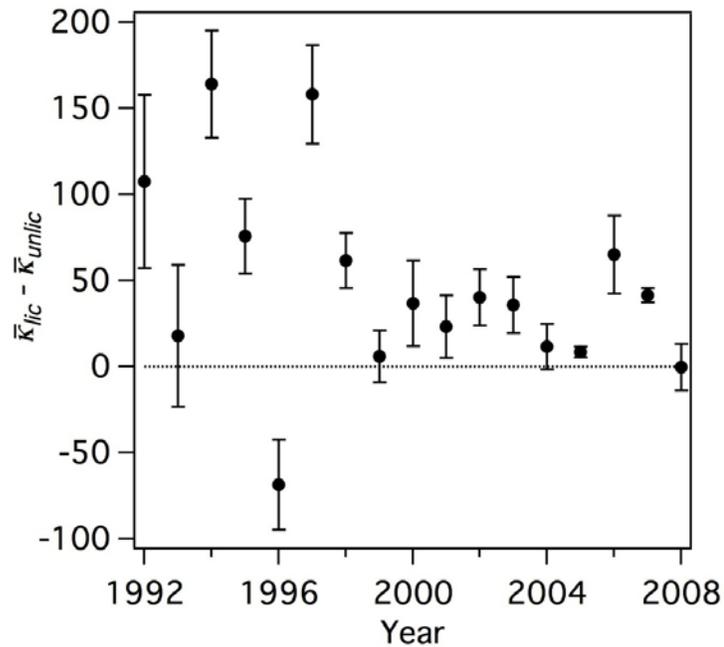


Figure 10. Predictive commercial value of patents using proposed Web hits relationship

4 Suggestions for Further Research

The data herein rely heavily upon a keyword identification of clean energy patents that focus on late-stage innovations. Future work needs to update and make available identification and categorization using a more general and flexible ontology comprised of a mixture of keyword, patent classification code, and assignee-based identification that is simultaneously inclusive of earlier-stage technologies but still accurate.

Significant work will need to be undertaken to further validate the Web-presence methodology. It would be extremely difficult to do this for the complete patent dataset due to the labor intensity of determining whether each patent has demonstrated commercial value. However, future work could include other institutions' licensed versus unlicensed patents as a method of calibrating this method. Qualitative surveys with active inventors will enrich this line of analysis. Lines of inquiry may include the inventor's qualitative assessment of the patent's value, the method of commercialization (start-up, license to large firm, license to non-profit, not deployed and why), and the level of commercial success. Additional focus on refining data quality around regional, organizational, collaborative, and institutional factors, as well as the funding sources for the research (particularly if the research was supported by a government grant) would be helpful. Also, further effort should be expended in the automated analysis of the content of each type of website. A variety of schemes could be investigated [87], including simple search term counts, dictionary look-up of appropriate terms, and semi-supervised machine learning algorithms (at least for the patents with enough hits apply such methods [88]).

We hope that this work will contribute to an open platform for future research by the community of science and policy scholars. The focus in this proposal on clean energy is just one slice of this effort. We hope to build, through the collaborative work of many scholars, a series of integrated databases that would enable the tracing of scientific investment from grants all the way to commercialization. Figure 11 illustrates an idealized schematic of this goal. This effort would help accelerate the development of a system where most grants could be traced to papers, papers to papers through citations, papers to patents through people and citations, and patents to commercial outcomes through people (for example, on boards of directors), licensing agreements, and the proposed Web-based measure of commercial impact and deployment. This would be a powerful tool for both DOE and the research community to better assess the value and success of research and development investments. This capability will be essential as the market quickly evolves, research investments increasingly come from the private sector, and there is a call for increased transparency on public research expenditures. Further development of this database and the addition of new datasets outside our control through an open-research model will enable this tool to be broadly used to investigate innovation, commercialization, and deployment for DOE and the public sector in general.

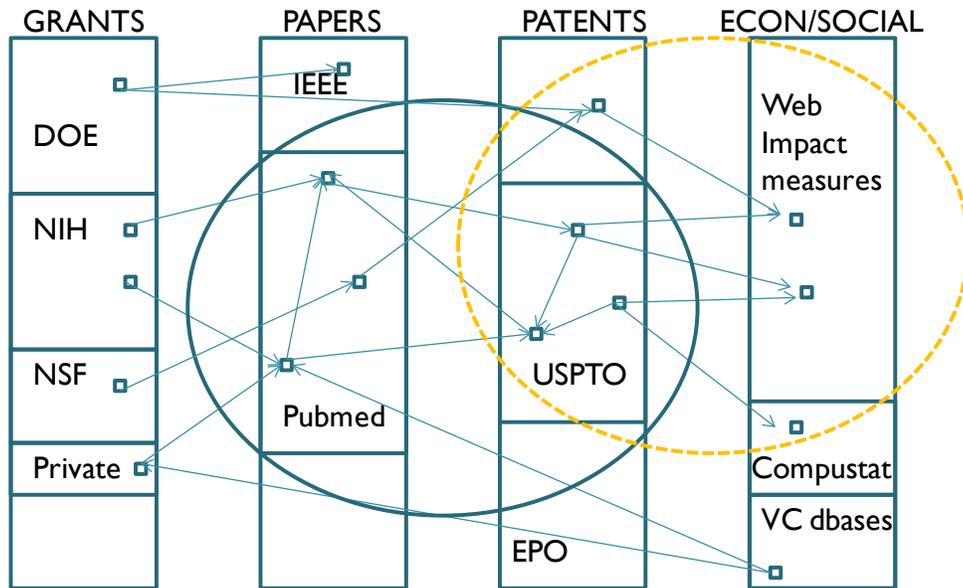


Figure 11. An idealized schematic of how future databases might be built and linked in order to understand how science investment influences economic outcomes

5 Conclusions

This work has expanded existing methods of patent analysis, introduced novel methods of Web presence analysis, and applied both to the empirical investigation of energy innovation. Preliminary results confirm and provide quantification of the notion that different energy technology areas exhibit unique patterns of invention and innovation. Generally, all energy sectors exhibit relatively isolated collaboration networks compared to the rich collaboration networks observed in other sectors (e.g., software and pharmaceuticals). The macro factors shaping these more collaborative industries help us understand the industrial, regulatory, and technical features that limit energy innovation and result in policy suggestions to drive richer collaboration in energy.

For energy as a whole, our results confirm the general notion that policymakers should encourage multidisciplinary collaboration, university and national laboratory commercialization efforts, and tailored support for regional innovation clusters. These insights may be extremely helpful in the development and implementation of two new initiatives, the Energy Frontier Research Centers and the Energy Innovation Hubs. These initiatives are focused on discrete technology questions with specific innovation targets. The analytical tools now at our disposal could help evaluate what are important social, sectoral, market, and geographic considerations. For example, the Hubs are envisioned as “multidisciplinary, multi-investigator, multi-institutional integrated research centers ... each comprise[d of] a large set of investigators spanning science, engineering, and policy disciplines focused on a single critical national need identified by the Department [of Energy] [89].” With enough data resolution, the analytical tools described in this work can help evaluate the proposed institutional arrangements and identify structures that can improve the likelihood of success.

Importantly, our results indicate that the dynamics of energy innovation are heterogeneous across sectors, and the relative benefits of a particular strategy are unlikely to translate smoothly from one technology area to another. Policies to foster innovation and the deployment of those innovations need to be tailored to each specific technology and its revealed pattern of innovation. This capability is important as the DOE embarks on a series of experiments to drive innovation across multiple sectors.

With further work, we expect to increase the size and interconnectedness of the energy patents database, achieve higher resolution on social, geographic, and institutional variables, and refine the statistical methods of investigating linkages between morphologies of invention and the production of breakthroughs. Together, these objectives will constitute a novel analytic and technical foundation for informing innovation policymaking.

References

1. Goldemberg, J. "Ethanol for a Sustainable Energy Future." *Sci.*; Vol. 315, Iss. 5813, 2007; pp. 808–810.
2. Li, X. "Diversification and Localization of Energy Systems for Sustainable Development and Energy Security." *Energy Policy*; Vol. 33, Iss. 17, 2005; pp. 2237–2243.
3. Wei, M.; Patadia, S.; Kammen, D. "Putting Renewables and Energy Efficiency to Work: How Many Jobs can the Clean Energy Industry Generate in the US?" *Energy Policy*; Vol. 38, Iss. 2, 2010; pp. 919–931.
4. Jaffe, A.; Peterson, S.R.; Portney, P.R.; Stavins R.N. "Environmental Regulation and the Competitiveness of U.S. Manufacturing: What Does the Evidence Tell Us?" *J. Econ. Lit.*; Vol. 33, Iss. 1, 1995; pp. 132–163.
5. Pacala, S.; Socolow, R. "Stabilization Wedges: Solving the Climate Problem for the Next 50 Years with Current Technologies." *Sci.*; Vol. 305, Iss. 5686, 2004, pp. 968-972.
6. Fthenakis, V.; Kim, H.C. "Life-cycle Uses of Water in U.S. Electricity Generation." *Renewable Sustainable Energy Rev.*; Vol. 14, Iss. 7, 2010; pp. 2039–2048.
7. Hansen, J.; Johnson, D.; Lacis, A.; Lebedeff, S.; Lee, P.; Rind, D.; Russell, G. "Climate Impact of Increasing Atmospheric Carbon Dioxide." *Sci.*; Vol. 213, Iss. 4511, 1981; pp. 957–966.
8. Thomas, C.D.; Cameron, A.; Green, R.E.; Bakkenes, M.; Beaumont, L.J.; Collingham, Y.C.; Erasmus, B.F.N.; Ferreira de Siqueira, M.; Grainger, A.; Hannah, L.; Hughes, L.; Huntley, B.; van Jaarsveld, A.S.; Midgley, G.F.; Miles, L.; Ortega-Huerta, M.A.; Peterson, A.T.; Phillips, O.L; Williams, S.E. "Extinction Risk from Climate Change." *Nat.*; Vol. 427, Iss. 6970, 2004; pp. 145–148.
9. Bernstein, L.; Bosch, P.; Canziani, O.; Chen, Z.; Christ, R.; Davidson, O.; Hare, W.; Huq, S.; Karoly, D.; Kattsov, V.; Kundzewicz, Z.; Liu, J.; Lohmann, U.; Manning, M.; Matsuno, T.; Menne, B.; Metz, B.; Mirza, M.; Nicholls, N.; Nurse, L.; Pachauri, R.; Palutikof, J.; Parry, M.; Qin, D.; Ravindranath, N.; Reisinger, A.; Ren, J.; Riahi, K.; Rosenzweig, C.; Rusticucci, M.; Schneider, S.; Sokona, Y.; Solomon, S.; Stott, P.; Stouffer, R.; Sugiyama, T.; Swart, R.; Tirpak, D.; Vogel, C.; Yoh, G. *Climate Change: Synthesis Report: An Assessment of the Intergovernmental Panel on Climate Change, IPCC Plenary XXVII (Valencia, Spain, 12-17 November 2007)*.
10. Stern, N. *The Stern review report on the economics of climate change*. Cambridge University Press, Cambridge, UK, 2006.
11. Anadon, L.; Bunn, M.; Chan, G.; Chan, M.; Gallagher, K.S.; Jones, C.; Kempener, R.; Lee, A.; Narayanamurti, V. *DOE FY 2011 Budget Request for Energy Research, Development, Demonstration, and Deployment: Analysis and Recommendations*, in *Report for Energy Technology Innovation Policy Research Group, Belfer Center for Science and International Affairs*. Cambridge, MA: Harvard University, Kennedy School of Government, 2010.

12. Kempener, R.; Anadon, L.D.; Tarco, J.C. *Energy Innovation Policy in Major Emerging Countries*. Cambridge, MA: Belfer Center for Science and International Affairs, Harvard University, 2010.
13. Pisano, G.P.; Shih, W.C. "Restoring American Competitiveness." *Harv. Bus. Rev.*; July-August 2009.
14. Porter, M.E.; Stern, S. "Innovation: Location Matters." *MIT Sloan Management Rev.*; Vol. 42, Iss. 4, 2001; pp. 28–36.
15. Jamison, E. *From Innovation to Infrastructure: Financing First Commercial Clean Energy Projects*. CalCEF, 2010, San Francisco, CA.
16. Rotman, D. "Cash for Infrastructure." *Technol. Rev.*, September/October 2010.
17. Tassef, G. "Rationales and Mechanisms for Revitalizing US Manufacturing R&D Strategies." *J. Technol. Transfer*; Vol. 35, 2010; pp. 283–333.
18. Margolis, R.M.; Kammen, D.M. "Evidence of Under-investment in Energy R&D in the United States and the Impact of Federal Policy." *Energy Pol.*; Vol. 27, Iss. 10, 1999; pp. 575–584.
19. Nemet, G.F.; Kammen, D.M. "U.S. Energy R&D: Declining Investment, Increasing Need, and the Feasibility of Expansion." *Energy Pol.*; Vol. 35, Iss. 1, 2007; pp. 746–755.
20. Margolis, R.M.; Kammen, D.M. "Underinvestment: The Energy Technology and R&D Policy Challenge." *Sci.*; Vol. 285, Iss. 5428, 1999; p. 690.
21. Biden, J. *The Recovery Act: Transforming the American Economy through Innovation*. 2010, The Whitehouse, Washington, D.C.
22. International Energy Agency. *Energy Technology Perspectives*. International Energy Agency, Paris, France, 2008.
23. Augustine, N.; Burns, U.; Doerr, J.; Gates, B.; Holliday, C.; Immelt, J.; Solso, T. *A Business Plan for America's Energy Future*. American Energy Innovatino Council, Washington, D.C., 2010.
24. Pachauri, R.K.; Reisinger, A. *Climate Change 2007: Synthesis Report*. 2007. IPCC, Geneva, Switzerland.
25. *Crossing the Valley of Death: Solutions to the Next Generation Clean Energy Project Financing Gap*. Bloomberg New Energy Finance, New York, NY, 2010.
26. Thomas, P.; Breitzman, A. *Patent Power Scorecards: Japan Ascendant*. New York, NY: IEEE Spectrum, March 2010.
27. Atkinson, R. "Expanding the R&E Tax Credit to Drive Innovation, Competitiveness and Prosperity." *J. Technol. Transfer*; Vol. 32, Iss. 6, 2007; pp. 617–628.
28. Heslin Rothenberg Farley & Mesiti P.C. *Clean Energy Patent Growth Index (CEPGI)*. 2010.
29. Singh, J.; Fleming, L. "Lone Inventors as Sources of Breakthroughs: Myth or Reality?" *Manage. Sci.*; Vol. 56, Iss. 1, 2010; pp. 41–56.
30. Gans, J.S.; Stern, S. "Is There a Market for Ideas?" *Ind. Corporate Change*; Vol. 19, Iss. 3, 2007; pp. 805–837.
31. Henderson, R.M.; Newell, R.G. *Accelerating Innovation In Energy: Insights from Multiple Sectors*. Boston, MA: H.B. School, 2010.

32. Acs, Z.J.; Anselin, L.; Varga, A. "Patents and Innovation Counts as Measures of Regional Production of New Knowledge." *Res. Pol.*; Vol. 31, Iss. 7, 2002; pp. 1069–1085.
33. Fleming, L. "Breakthroughs and the 'Long Tail' of Innovation." *MIT Sloan Manage. Rev.*; Vol. 49, 2007; pp. 69–77.
34. Trajtenberg, M. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *Rand J. Econ.*; Vol. 21, Iss. 1, 1990; pp. 172–187.
35. Lai, R.; D'Amour, A.; Fleming, L. *The Careers and Co-authorship Networks of U.S. Patent-holders, Since 1975*. Cambridge, MA: Harvard Business School, Harvard Institute for Quantitative Social Science, 2009.
36. Cardona, V.A. *Clean Energy Patent Growth Index (CEPGI)*. Albany, NY: H.R.F.M. P.C., 2009.
37. Fleming, L.; Mingo, S.; Chen, D. "Collaborative Brokerage, Generative Creativity, and Creative Success." *Admin. Sci. Q.*; Vol. 52, Iss. 3, 2007; pp. 443–475.
38. Popp, D.; Juhl, T.; Johnson, D.K.N. "Time In Purgatory: Examining the Grant Lag for U.S. Patent Applications." *Top. Econ. Anal. Pol.* Vol. 4, Iss. 1, 2004.
39. Rosvall, M.; Bergstrom, C.T. "Maps of Random Walks on Complex Networks Reveal Community Structure." *Proc. Natl. Acad. Sci.*; Vol. 105, Iss. 4, 2008; pp. 1118–1123.
40. Sorenson, A.; Fleming, L. "Science and the Diffusion of Knowledge." *Res. Pol.*; Vol. 33, Iss. 10, 2004; pp. 1615–1634.
41. Alcacer, J.; Gittelman, M. "How Do I Know What You Know? Patent Examiners and the Generation of Patent Citations." *Rev. Econ. Stat.*; Vol. 88, Iss. 4, 2006; pp. 774–779.
42. Hall, B.; Jaffe, A.; Trajtenberg, M. "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools." NBER Working Paper No. 8498, 2001. October.
43. Bonzi, S.; Snyder, H.W. "Motivations for Citation: A Comparison of Self Citation and Citation to Others." *Scientometrics*; Vol. 21, Iss. 2, 1991; pp. 245–254.
44. Aksnes, D.W. "A Macro Study of Self-citation." *Scientometrics*; Vol. 56, Iss. 2, 2003; pp. 235–246.
45. West, J.; Althouse, B.; Rosvall, M.; Bergstrom, C.; Bergstrom, T. *Methods Version 2.01, The Eigenfactor™ Metrics*. 2008.
46. Fowler, J.H.; Aksnes, D.W. "Does Self-citation Pay?" *Scientometrics*; Vol. 72, Iss. 3, 2007; pp. 427–437.
47. Podolny, J.M.; Stuart, T.; Hannan, M.T. "Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984-1991." *AJS*; Vol. 102, Iss. 3, 1996; pp. 659–689.
48. Thompson, P.; Fox-Kean, M. "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment." *Amer. Econ. Rev.*; Vol. 95, Iss. 1, 2005; pp. 450–460.
49. Scherer, F.M.; Harhoff, D. "Policy Implications for a World with Skew-Distributed Returns to Innovation." *Res. Pol.*; Vol. 29, 2000; pp. 559–566.
50. MacRoberts, M.H.; MacRobert, B.R. "Problems of Citation Analysis." *Scientometrics*; Vol. 36, Iss. 3, 1996; pp. 435–444.

51. Hall, B.H. "Innovation and Market Value." *NBER Working Paper Series*, 1999.
52. White, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*; Vol. 48, Iss. 4, 1980; pp. 817–838.
53. Wooldridge, J. "Distribution-free Estimation of Some Nonlinear Panel Data Models." *J. Econometrics*; Vol. 90, Iss. 1, 1999; pp. 77–97.
54. Long, J. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications, 1997.
55. Zellner, C. "The Economic Effects of Basic Research: Evidence for Embodied Knowledge Transfer Via Scientists' Migration." *Res. Pol.*; Vol. 32, Iss. 10, 2003; pp. 1881–1895.
56. King, G.; Tomz, M.; Wittenberg, J. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *Amer. J. Polit. Sci.*; Vol. 44, Iss. 2, 2000; pp. 347–361.
57. Tomz, M.; Whittenberg, J.; King, G. "Clarify: Software for Interpreting and Presenting Statistical Results." *J. Stat. Software*; Vol. 8, Iss. 1, 2003; pp. 1–30.
58. Zellner, B.A. "Using Simulation to Interpret Results from Logit, Probit, and Other Nonlinear Models." *Strategic Manage. J.*; Vol. 30, Iss. 12, 2009; pp. 1335–1348.
59. Davis, M. "Contrast Coding in Multiple Regression Analysis: Strengths, Weaknesses, and Utility of Popular Coding Structures." *J. Data Sci.*; Vol. 8, 2010; pp. 61–73.
60. Lewis, J. "Technology Acquisition and Innovation in the Developing World: Wind Turbine Development in China and India." *Stud. Comp. Int. Dev.*; Vol. 42, Iss. 3–4, 2007; pp. 208–232.
61. Levi, M.A.; Economy, E.C.; O’Neil, S.K.; Segal, A. *Energy Innovation: Driving Technology Competition and Cooperation Among the U.S., China, India, and Brazil*. 2010, Council on Foreign Relations: New York.
62. Tijssen, R.J.W. "Global and Domestic Utilization of Industrial Relevant Science: Patent Citation Analysis Of Science–Technology Interactions and Knowledge Flows." *Res. Pol.*; Vol. 30, Iss. 1, 2001; pp. 35–54.
63. Deutch, J.M.; Forsberg, C.W.; Kadak, A.C.; Kazimi, M.S.; Moniz, E.J.; Parsons, J.E.; Du, Y.; Pierpoint, L. "Update of the MIT 2003 Future of Nuclear Power." 2009, Massachusetts Institute of Technology: Cambridge, MA.
64. Kidd, S. "New Nuclear Build – Sufficient Supply Capability?" *Nucl. Eng. Int.*; Vol. 54, Iss. 656, 2009; pp. 10–11.
65. Lane, J. "Assessing the Impact of Science Funding." *Sci.*; Vol. 324, Iss. 5932, 2009; pp. 1273–1275.
66. Lane, J. "Let's Make Science Metrics More Scientific." *Nat.*; Vol. 464, Iss. 7288, 2010; pp. 488–489.
67. Wuchty, S.; Jones, B.F.; Uzzi, B. "The Increasing Dominance of Teams in Production of Knowledge." *Sci.*; Vol. 316, Iss. 5827, 2007; pp. 1036–1039.
68. Harhoff, D.; Narin, F.; Scherer, F. M.; Vopel, K. "Citation Frequency And The Value Of Patented Inventions." *Rev. Econ. Stat.*; Vol. 81, Iss. 3, 1999; pp. 511–515.
69. Torvik, V.; Smalheiser, N. "Author Name Disambiguation in MEDLINE." *ACM Trans. Knowl. Discovery Data*; Vol. 3, Iss. 11, 2009; pp. 1–29.

70. Bonvillian, W.; Weiss, C. *Structuring an Energy Technology Revolution*. Cambridge, MA: MIT Press, 2009.
71. Jacobsson, S.; Bergek, A. "Transforming the Energy Sector: The Evolution of Technological Systems in Renewable Energy Technology." *Ind. Corporate Change*; Vol. 13, Iss. 5, 2004; pp. 815–849.
72. Holdren, J. "The Energy Innovation Imperative: Addressing Oil Dependence, Climate Change, and Other 21st Century Energy Challenges." *Innovations*; Vol. 1, Iss. 2, 2006; pp. 3–23.
73. Smil, V., *Energy at the Crossroads: Global Perspectives and Uncertainties*. Cambridge, MA: MIT Press, 2003.
74. Coombs, R.; Narandren, P.; Richards, A. "A Literature-based Innovation Output Indicator." *Res. Pol.*; Vol. 25, Iss. 3, 1996; pp. 403–413.
75. Giles, J. "Internet Encyclopaedias Go Head to Head." *Nat.* Vol. 438, Iss. 900–901, 2005.
76. Breitzman, A.; Thomas, P. "Using Patent Citation Analysis to Target/Value MA Candidates." *Res.-Technol. Manage.*; Vol. 45, 2002; pp. 28–36.
77. Karlenzig, W. "Top Five U.S. Cities for Cleantech." <http://www.sustainlane.us/articles/cleantech.jsp>. Accessed January 4, 2011.
78. Fleming, L. "Research - Perfecting Cross-pollination." *Harv. Bus. Rev.*; Vol. 82, Iss. 9, 2004; pp. 22–24.
79. Johnson, S.B. *Where Good Ideas Come From: The Natural History of Innovation*. New York, NY: Riverhead Books, 2010.
80. Hidalgo, C.; Klinger, B.; Barabasi, A.-L.; Hausmann, R. "The Product Space Conditions the Development of Nations." *Sci.* Vol. 317, Iss. 482–487, 2007.
81. Hidalgo, C.; Hausmann, R. "The Building Blocks of Economic Complexity." *Proc. Natl. Acad. Sci.*; Vol. 106, Iss. 26, 2009; pp. 10570–10575.
82. Cooke, P. "Regional Innovation Systems, Clusters, and the Knowledge Economy." *Ind. Corporate Change*; Vol. 10, Iss. 4, 2001; pp. 945–974.
83. Asheim, B.; Isaksen, A. "Regional Innovation Systems: The Integration of Local 'Sticky' and Global 'Ubiquitous' Knowledge." *J. Technol. Transfer*; Vol. 27, Iss. 1, 2002; pp. 77–86.
84. Hamel, G. "Bringing Silicon Valley Inside." *Harv. Bus. Rev.*; Vol. 77, Iss. 5, 1999; pp. 70–84.
85. Lécuyer, C. *Making Silicon Valley: Innovation and the Growth of High Tech, 1930-1970*. Cambridge, MA: MIT Press, 2006.
86. Saxenian, A. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press, 1994.
87. Krippendorff, K. *Content Analysis: An Introduction to its Methodology*. Thousand Oaks, CA: S. Publications, 2004.
88. Hopkins, D.; King, G. "A Method of Automated Nonparametric Content Analysis for Social Science." *Amer. J. Political Sci.*; Vol. 54, Iss. 1, 2010; pp. 229–247.
89. Energy.gov. "Q&A." <http://www.energy.gov/hubs/qanda.htm>. Accessed January 12, 2010.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Executive Services and Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

| | | | | | | |
|---|------------------------------------|-------------------------------------|---|---|--|--|
| 1. REPORT DATE (DD-MM-YYYY) March 2011 | | | 2. REPORT TYPE Technical Report | | 3. DATES COVERED (From - To) | |
| 4. TITLE AND SUBTITLE Clean Energy Innovation: Sources of Technical and Commercial Breakthroughs | | | | 5a. CONTRACT NUMBER DE-AC36-08GO28308 | | |
| | | | | 5b. GRANT NUMBER | | |
| | | | | 5c. PROGRAM ELEMENT NUMBER | | |
| 6. AUTHOR(S) Thomas D. Perry IV and Mackay Miller (NREL) Lee Fleming (Harvard Business School) Kenneth Younge (University of Colorado) James Newcomb (NREL, Current Affiliation: Rocky Mountain Institute) | | | | 5d. PROJECT NUMBER NREL/TP-6A20-50624 | | |
| | | | | 5e. TASK NUMBER SA10.102C | | |
| | | | | 5f. WORK UNIT NUMBER | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Renewable Energy Laboratory 1617 Cole Blvd. Golden, CO 80401-3393 | | | | 8. PERFORMING ORGANIZATION REPORT NUMBER NREL/TP-6A20-50624 | | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | | | 10. SPONSOR/MONITOR'S ACRONYM(S) NREL | | |
| | | | | 11. SPONSORING/MONITORING AGENCY REPORT NUMBER | | |
| 12. DISTRIBUTION AVAILABILITY STATEMENT National Technical Information Service U.S. Department of Commerce 5285 Port Royal Road Springfield, VA 22161 | | | | | | |
| 13. SUPPLEMENTARY NOTES | | | | | | |
| 14. ABSTRACT (Maximum 200 Words) Low-carbon energy innovation is essential to combat climate change, promote economic competitiveness, and achieve energy security. Using U.S. patent data and additional patent-relevant data collected from the Internet, we map the landscape of low-carbon energy innovation in the United States since 1975. We isolate 10,603 renewable and 10,442 traditional energy patents and develop a database that characterizes proxy measures for technical and commercial impact, as measured by patent citations and Web presence, respectively. Regression models and multivariate simulations are used to compare the social, institutional, and geographic drivers of breakthrough clean energy innovation. Results indicate statistically significant effects of social, institutional, and geographic variables on technical and commercial impacts of patents and unique innovation trends between different energy technologies. We observe important differences between patent citations and Web presence of licensed and unlicensed patents, indicating the potential utility of using screened Web hits as a measure of commercial importance. We offer hypotheses for these revealed differences and suggest a research agenda with which to test these hypotheses. These preliminary findings indicate that leveraging empirical insights to better target research expenditures would augment the speed and scale of innovation and deployment of clean energy technologies. | | | | | | |
| 15. SUBJECT TERMS patents; clean energy; Web hits; innovation; breakthroughs; low carbon | | | | | | |
| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT UL | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON | |
| a. REPORT Unclassified | b. ABSTRACT Unclassified | c. THIS PAGE Unclassified | | | 19b. TELEPHONE NUMBER (Include area code) | |