NBER WORKING PAPER SERIES

USING SCIENTIFIC PUBLICATIONS TO EVALUATE GOVERNMENT R&D SPENDING: THE CASE OF ENERGY

David Popp

Working Paper 21415
http://www.nber.org/papers/w21415

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2015

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

Using Scientific Publications to Evaluate Government R&D Spending: The Case of Energy
David Popp
NBER Working Paper No. 21415
July 2015
JEL No. O21,O38,Q42,Q48,Q55

ABSTRACT

The mix of public and private research funding investments in alternative energy presents a challenge for isolating the effect of government R&D funding. Factors such as energy prices and environmental policy influence both private and public R&D decisions. Moreover, because government R&D is further upstream from the final commercialized product, it may take several years for its effect on technology to be realized. Combining data on scientific publications for alternative energy technologies with data on government R&D support for these technologies, we address these challenges. First, we ask how long it takes for energy R&D to provide successful research outcomes. We both provide information on the lags between research funding and new publication and link these articles to citations in U.S. energy patents. One million dollars in additional government R&D funding leads to 1-2 additional publications, but with lags as long as ten years between initial funding and publication. Second, we ask whether adjustment costs associated with large increases in research funding result in diminishing returns to government R&D. There is no evidence of diminishing returns on the level of publication output, but some evidence that additional funding leads to lower quality publications, using citations as a measure of publication quality.

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

I. Introduction

Because of the long-term, uncertain nature of climate change, government R&D funding plays an important role in long-term strategies to reduce greenhouse gas emissions. The mix of public and private research funding for technologies to reduce emissions presents a challenge for estimating the effect of government-funded R&D. Because it is further upstream from the final commercialized product, government R&D should take longer to have an observable effect on outcomes than private R&D. At the same time, both private and public R&D are driven by the same demand-side influences, such as energy prices and environmental policy, making it difficult to disentangle the effect of each.

Such funding is particularly important for alternative energy sources, many of which are still too costly to be competitive with fossil fuels without policy support. Generation of electricity and heat is the largest source of carbon emissions, accounting for 42% of carbon emissions worldwide in 2012 (IEA 2014). Meeting the climate policy goals currently under consideration, such as European Union discussions to reduce emissions by 40 percent below 1990 levels by 2030, will not be possible without replacing much of this electric generating capacity with alternative, carbon-free energy sources.[1]

In this paper, we use data on scientific publications to assess the effect of government-sponsored energy R&D. While many economists now use patent data to evaluate energy R&D (Popp 2001, 2002, 2006; Johnstone *et al*. 2010 Verdolini and Gaelotti, 2011), publication data are used less frequently within the field. For evaluating public research funding efforts, publication data provide a more appropriate outcome measure than patents. Moreover, while direct evidence on the relationship between public R&D funding and energy costs is desirable, the long lags

---

[1] These limits come from http://ec.europa.eu/clima/policies/international/negotiations/future/index_en.htm, accessed January 7, 2015.

between basic research support and commercial outcomes make measuring the final impact of basic research difficult (Lane, 2009). For example, the costs of renewable energy sources depend not only on knowledge accumulated from public R&D, but also on advances due to private R&D, both of which are driven by many of the same demand signals. By looking at the effect of public R&D funding on publications, this paper attempts to isolate the effect of public R&D to shed light on the process through which public R&D helps develop scientific knowledge. In addition, publication data are available for a wide range of countries and technologies. In contrast, studies using direct measures of energy research outcomes, such as changes in renewable energy costs, use data from only a few countries (e.g. Söderholm and Klaasen 2007).

However, compared to patents, publication data bring additional challenges to evaluate properly. Scientific publications are observed for papers that have been accepted by a journal editor. Editors not only judge the potential article for quality, but also must consider whether it will be of interest to readers. Thus, our econometric model accounts for factors influencing both the willingness of scientists to study renewable energy and the willingness of editors to publish such articles. Using this framework, we combine data on scientific publications for alternative energy technologies, such as wind and solar power, with data on government R&D support for specific energy technologies and controls such as energy prices and the nature of electricity production in each country. Such controls are important to assess the marginal impact of additional government R&D. We use these data to address two research questions.

First, we ask how long it takes for energy R&D to provide successful research outcomes. Within the energy R&D literature, most studies consider only a single year of energy R&D or construct aggregated stocks of past R&D efforts using predetermined rates of decay. By focusing on both inputs and outputs of the research process, this study provides new evidence on the timing of these flows and suggests that simply focusing on a single year of energy R&D omits longer run effects from R&D spending. Using both finite distributed lag and polynomial distributed lag models, we provide information on the lags between research funding and new publications. Then, we link these articles to citations in U.S. energy patents. Asking how long it takes for publications to be cited by a patent helps illuminate the lags between basic science and applied work.

Second, we ask whether adjustment costs associated with large increases in research funding result in diminishing returns to government R&D. Despite concerns raised about the adjustment costs of dramatic changes in funding for medical science (Freeman and van Reenen 2009), Schulke-Leech (2014) provides evidence of similar experiences for energy R&D budgets, raising the concern that such volatility may be particularly problematic for energy technology research, which requires large capital investment. As scientists' time is limited and the supply of researchers able and willing to work on a given topic is inelastic in the short run, large increases in funding may not lead to corresponding increases in scientific output (e.g. Goolsbee 1998). Moreover, assuming that funding agencies support the highest quality projects first, as funding increases and governments provide support to a greater number of projects, the additional projects supported may be of lower quality than those supported in leaner budget years. In this paper, we use counts of publications to ask whether there are diminishing returns to the quantity of publications produced from new energy R&D funding, and citation data to ask whether the quality of publications fall as research output increases.

II. Literature Review

Within the field of economics, there is a long tradition of evaluating the research process using econometric techniques that provide a theoretical framework for analysis. However, until recently, most of this analysis has used patent data, rather than publication data, as the measure of research output.[2] While the use of publication data within the economic research community has recently increased, the focus has primarily been on using publication data to measure knowledge flows, rather than evaluating the effectiveness of research inputs.[3] Research using publication data to examine the returns to research includes Jacob and Lefgren (2011), who find that receiving an National Institutes of Health (NIH) NIH grant leads to just one additional publication over five years. Jacob and Lefgren suggest that the small effect of NIH funding comes from non-recipients simply switching to other sources of funding. The potential for other sources of funding to serve as substitutes raises questions about the true marginal effects of government research funds. Using university level data, Rosenbloom *et al.* (2014) find a larger return on public chemistry R&D funding at U.S. universities. However, they do not provide estimates of the lag structure of R&D. Showing the importance of other demand factors, Bhattacharya and Pachalen (2011) use publication data to examine medical research across related technological areas. They link scientific opportunity, measured by the availability of new drugs, to scientific publications referencing both these ingredients and these medicines. While they find that both research opportunities and potential market size matter, they do not directly evaluate the effects of research inputs, such as public R&D spending.

Within the energy innovation literature, most studies evaluating the effect of publicly funded energy R&D focus on the effect of public R&D on new energy patents or on the cost of

[2] Early exceptions include Pardey (1989), who finds that long-term trends in agricultural R&D spending matter more than short-run, year-to-year variation within states; Adams (1990), who uses publications as an explanatory variable, linking scientific knowledge, as represented by an accumulated stock of publications, to output growth; and Adams and Griliches (1996), who estimate a production function of knowledge produced by 110 top U.S. research universities, finding a positive relationship between R&D and publications, but an even stronger link between employment of scientists and engineers at universities and publication counts.
[3] Examples include Adams, Clemmons, and Stephan (2004, 2006), Adams and Clemmons (2008), Jones *et al*. (2008), Wuchty *et al*. (2007), Azoulay, Zivin, and Wang (2010), Zucker *et al*. (2007), and Furman *et al*. (2012) .

alternative energy sources. Looking at patents, Popp (2002) uses a distributed lag framework and finds a limited role for government R&D, with government energy R&D serving as a substitute for private energy R&D during the 1970s, but as a complement to private energy R&D afterwards. Other papers consider just a single year of public R&D data, either contemporary or lagged (e.g. Johnstone *et al* 2010, Verdolini and Gaelotti 2011, Dechezleprêtre and Glachant 2014, Nesta *et al*. 2014). While these studies typically find a positive effect of public R&D on patenting, the short lags raise questions about what is truly being identified.

Similarly, a second line of research explores the role of both R&D investment and experience using two-factor learning curves, modeling cost reductions as a function of both cumulative capacity (learning-by-doing, or LBD) and R&D (learning-by-searching, or LBS). Examples include Klaasen *et al.* (2005), Söderholm and Sundqvist (2007), Söderholm and Klaasen (2007) and Ek and Söderholm (2010). To be comparable with the notion of cumulative capacity, R&D is typically aggregated into a stock of R&D capital.[4] Thus, endogeneity is a concern, as both investments in capacity and past R&D expenditures are simultaneously determined. Ek and Söderholm model R&D choice directly, treating public R&D as a function of the real investment cost of wind, the opportunity cost of public R&D, measured by real rate of return on long-term treasury bonds, and the share of total government debt. They estimate a learning-by-searching rate of about 20 percent, but find it is only significant at the 12 percent level.

As these papers show, our understanding of the timing of energy R&D's potential benefits remains limited. Most studies consider only a single year of energy R&D or construct aggregated stocks of past R&D efforts using predetermined rates of decay. Related research in other fields offers some evidence, but remains incomplete. Research on the productivity of publicly funded

[4] Söderholm's work typically lags the R&D stock by two-years and assumes a depreciation rate of three percent.

health R&D illustrates the challenges of identifying the effect of public R&D funding that comes at the beginning of a long, uncertain research process. In medical research, the discovery of a potential new drug is followed by years of clinical studies and human testing. The average lag between the beginning of human testing to FDA application is nine years, and R&D to explore potential new drugs occurs years before testing can begin. Toole (2012) studies the effect of public basic research funding from the U.S. National Institutes of Health (NIH) on applications for new molecular entities (NMEs) approved by the Food and Drug Administration (FDA). To make the regressions manageable given the long lags necessary to develop a new treatment, Toole creates separate stocks of accumulated private and public R&D. His results confirm that the lag between public R&D funding and new product development can be large, estimating a lag between initial public R&D investment and NME application between 17 and 24 years. Similarly, Blume-Kohout (2012) examines the links between NIH funding and the number of drugs entering clinical testing for a range of diseases, controlling for potential market size. Unlike Toole, she uses a finite distributed lag model, including up to 12 years of lagged R&D funding. Because of multicollinearity concerns, she focuses on the aggregate long-run effect of NIH R&D funding, with a 10 percent increase in NIH funding leading to a 4.5 percent increase in clinical trials after 12 years, but does not provide information on the how the effect of R&D from a particular year varies over time.

Using macro-level data, Crespi and Geuna (2008) find an optimal lag length between 6 and 7 years in their study of higher education research and development spending on publications across 14 OECD countries from 1981-2002. However, their study includes research on a wide range of topics, and does not consider differences among fields, nor other demand factors that may influence scientific research. In contrast, in this paper I both provide estimates of the year-by-year effect of energy R&D on publications, similar to Crespi and Geuna, while also controlling for other market-driven and policy forces that affect the direction of basic research. This allows me to assess the extent to which increases to public energy R&D funding *within a specific field* lead to increased scientific output *within that field*.

III. Estimation

We estimate the effects of public R&D on scientific publications for each of four technologies: biofuels, energy efficiency, solar energy, and wind energy. While many economists now use patent data to evaluate R&D, publication data are used less frequently within the field for this purpose. Both patents and publications are the *outcome* of a research process. To observe either, the researcher must have a successful project and the appropriate authority (patent examiner or journal editor) must decide to accept the completed project to patent or publish. For patents, the only threshold is a legal threshold. In contrast, the acceptance process for a research paper is more complex. The number of publications observed in a given year is the equilibrium outcome of a demand process, in which editors choose desired articles based on the perceived importance to their readers, and a supply process, in which researchers produce articles to submit in response to incentives provided by research funding and perceived prestige. Editors not only judge the potential article for quality, but also must consider whether it will be of interest to readers. This complicates analyzing the effect of public R&D spending on publications, as many of the same factors determining public R&D decisions likely affect reader interest, and thus the editor's decision to publish

Controlling for other factors influencing research interest in a given technology is also important, as it affects the supply of articles available to publish. For instance, even researchers without public funding are likely to find alternative energy an attractive area of research when energy prices are high. While we do not observe alternative funding sources that may be available to researchers, we do observe other factors affecting the availability of these funds, such as energy prices and policy. Omitting these variables erroneously assumes that all new research is the result of government R&D funding. However, controlling for such factors complicates the model, as the same factors that provide scientists incentives to work on alternative energy may also influence government research funding decisions.

We use two alternative approaches to address this concern. First, the regressions include both country and year fixed effects. Many of the factors influencing *both* R&D spending and publication on a specific technology will be country-specific. For example, countries with abundant sunshine should both be willing to support solar R&D and have researchers actively publishing on solar energy. Similarly, year fixed effects account for both changes in the research opportunities available to scientists at a given time, which fluctuate as advances in science make some areas of research look more or less fruitful, and for the competing submissions from other fields that editors may have to choose from at any given time.

Second, to control for any remaining omitted variable bias that does vary over time, we also use instrumental variables for contemporaneous energy R&D. Potential instruments must be correlated with research spending, but not the publication decision. Thus, we include instruments for R&D spending on related technologies (e.g. using biofuels R&D as an instrument for solar energy R&D spending), as well as instruments that model the political process determining R&D funding: tax revenues (excluding social security) as a percentage of GDP, general government expenditure as a percentage of GDP, and a set of dummy variables representing the political leanings of the government. Specific instruments for each technology are discussed in section IV.

Because the research process takes place over time, lagged effects for the variables described above will be important. Articles are first observed on their publication date, *t*, and thus contribute to publication counts in this year. However, publication occurs at the end of a long process.[5] Indexing countries by *i*, we estimate the following model:

(1) $Q_{i,t} = \sum_{s=0}^{T} \beta_{t-s} R_{i,t-s} + \sum_{s=0}^{T} \boldsymbol{\gamma}_{\mathbf{t-s}} \mathbf{POLICY}_{i,t-s} + \sum_{s=0}^{T} \boldsymbol{\delta}_{\mathbf{t-s}} \mathbf{X}_{i,t-s} + \alpha_i + \eta_t + \epsilon_{i,t}$

In equation (1), $Q_{i,t}$ represents the number of publications from authors in country *i* published in year *t* and $R_{i,t-s}$ represents government R&D spending on a given technology by country *i* in year *t-s*. We separate the variables measuring interest in each energy source into two parts: $\mathbf{POLICY_{i,t-s}}$ represents policies relevant to the technology and $\mathbf{X_{i,t-s}}$ represents various control variables that affect demand for alternative energy sources in each country, such per capita GDP and the share of energy coming from hydropower and nuclear. Both the policy variables and other controls used vary by technology, and are discussed in the next section. α_i and η_t, represent the country and year fixed effects described above.

We use first-differenced panel data techniques to estimate equation (1). First differencing has two advantages over a fixed effect model for this estimation. Most importantly, first differencing avoids the problem of spurious regressions in the case where explanatory variables have a unit root (Wooldridge, 2012). I find evidence of unit roots for R&D using the IPS test (Im, Pesaran and Shin, 2003) both with and without a trend included. In all cases, I reject the null hypothesis of a unit root when using first-differenced data. In addition, first differencing does not depend on strict exogeneity of the explanatory variables. Strict exogeneity will not hold in the case of R&D funding if, for example, a positive shock to R&D productivity leads to more funding

[5] In the patent literature, researchers traditionally use the patent application date to avoid delays due to the examination process. Unfortunately, that is not possible with publication data, as the submission date is not known, and articles may need to be submitted at multiple journals before finding one to accept it.

in future years. In contrast, the first difference model only requires that $E[\Delta \mathbf{X}_{i,t}\, \Delta \varepsilon_{i,t}] = 0$ (Cameron and Trivedi, 2009). By using instruments for contemporaneous R&D, I am able to test whether this assumption holds and avoid bias by using instrumental variables for $\Delta RD_{i,t}$ if necessary. The instruments used vary by technology and are discussed in the next section.

Finally, because of multicollinearity concerns when using multiple lagged variables, I also estimate equation (1) using a polynomial distributed lag (PDL) model (e.g. Almon 1965). Rather than impose structure on the lag process, the PDL model uses polynomials of various degrees to proxy for the effect of the lagged variables. I retain the use of first-differenced variables and instruments for $\Delta RD_{i,t}$ in the PDL model.

IV. Data

The publication data come from the Thomson Reuters Web of Science database. Using a series of keyword searches of article titles, abstracts, and keywords, provided in Appendix A, we identified journal publications for each of our technologies. We focus on publications in scientific journals by dropping articles such as reviews, editorials, or news items. We do include proceedings papers that are included in the Web of Science database. The publication data run from 1991-2011, as complete records of titles, abstracts, and keywords begin in 1991. Once we identified appropriate keywords, Thomson Reuters provided a custom database containing all publications from 1991-2011 for our four technologies. The database includes descriptive information on each paper, including the date of publication and addresses for each author, which we use to assign articles to each country. For each technology, articles are aggregated by year and country. In the case of articles with multiple authors from multiple countries, we use weighted counts, assigning articles proportionately by the number of countries represented. For example, an article with 2 US authors and 1 Japanese author counts as 0.67 for the US and 0.33 for Japan.[6] In addition, the database also includes descriptive data on each article citing these energy articles, which we use as a proxy for article quality in section VE.

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

Government R&D data by country and technology are taken from the International Energy Agency (IEA). IEA data include technology-specific government energy technology R&D budgets for 26 IEA member countries. While annual data are available, the time series are incomplete for some countries. As such, the available R&D limit the number of countries included in the analysis. R&D data go back as far as 1974 for some countries. Using a starting date of 1992 for the publication data, we select countries with at least 10 years of lagged R&D data for a given

[6] We also ran models using unweighted counts, where we assign a full article to each country represented. The results do not change. It is not possible to assign a primary country by identifying the first or last author listed, as the order of addresses and authors in the database are not linked.

technology. Table 1 lists the countries included in the regressions for each technology.[7]

Table 2 provides descriptive statistics for both R&D and the weighted publications. R&D data are shown in millions of 2010 dollars. Evaluated at the means, for most technologies we see a bit less than one million dollars of R&D spending per publication. The exception is energy efficiency, for which a little over $3 million of R&D is spent per publication. While this may indicate that energy efficiency R&D is less productive than other R&D spending, it may also indicate that our keyword searches identify a lower percentage of relevant energy efficiency articles than for other technologies. Table 3 shows the top 10 sources of publications for each technology. While our sample includes most of the top publishing countries, note that some emerging economies such as China and Brazil are also actively publishing on alternative energy.

Figure 1 shows the aggregate trends in both energy R&D and publications for all countries in each regression. Not surprisingly, both R&D and publications have been increasing over time. The large increase in energy R&D in 2009 is primarily a result of the US stimulus spending and is a one-time shock to R&D. Figures 2 and 3 illustrate these trends for each country for solar and wind energy. Note that there is variation both across times and across countries, which is important for separately identifying the effect of energy R&D, as opposed to simply finding that publications increase due to increased global attention to climate change. For example, prior to the 2009 stimulus, the U.S. experiences two peaks in wind energy R&D – one after the 1970s energy crises, and another in the mid-1990s. In contrast, Denmark and the Netherlands have relatively flat R&D budgets, with the exception of the Netherlands in 1985. Note also that countries also choose to

[7] One complication with the R&D data is that the reported national data of European countries omits funding from the European Commission (IEA 2011). Fortunately, such funding represents a small percentage of overall public energy funding in Europe (Dechezleprêtre and Popp, 2015). Nonetheless, to insure that missing European-wide funding efforts do not affect our results, we include EU-specific year effects in the regressions to account for any EU-wide changes in R&D funding. These year effects have little effect on the final results, as shown in Appendix B, which compares the results with and without these EU-specific year fixed effects.

emphasize different technologies. For example, in 2010 the United States spent over ten times as much on wind energy as Japan in 2007, but nearly twice as much over twice as much on biofuels research as on solar energy, whereas Japan spent more on solar energy R&D than on biofuels. These figures also show the importance of lagged R&D effects, as within individual countries, R&D spending generally peaks several years before publication counts. Examples include wind energy in Germany, Italy, and Japan.

Table 4 lists the policy variables, controls, and instruments used to estimate equation (1) for each technology. The columns indicate which variables are included for each technology. Policies include gasoline taxes (for biofuels) and the level of renewable energy mandates and feed-in tariffs used to promote wind or solar energy.[8] In addition to per capita GDP, which is included in all regressions, other controls include gasoline prices for biofuels and energy efficiency. We do not include electricity prices in the model for solar or wind energy, as increased usage of solar and wind would lead to higher energy prices. Instead, for solar and wind we consider other factors influencing its demand, including the prevalence of clean energy substitutes (e.g. hydropower and nuclear) and the growth rate of electricity consumption. The extent to which a country needs such policies to reduce emissions depends on the carbon-intensity of electricity generation. Countries already making extensive use of other carbon-free energy sources, such as hydro or nuclear power, are expected to make less effort to promote renewable energy, and thus have lower interest in research on these technologies.[9] Similarly, when electricity consumption is growing more quickly, more investment in new electricity capacity will be needed, making investments in renewables more likely.

---

[8] Continuous variables are constructed for selected policy instruments based on the OECD Renewable Energy Policy Database (OECD-EPAU 2013) – an update and extension of the dataset originally used in Johnstone *et al*. (2010).
[9] For instance, Popp *et al*. (2011) shows that investment in renewable energy capacity is lower in countries using more nuclear or hydro power.

As noted earlier, our instruments include two categories of variables. We include instruments for R&D spending on related technologies, with the specific technologies chosen for each technology indicated in the table.[10] Overidentification tests were used to verify the validity of all technologies included and to rule out invalid instruments. In addition, since all energy R&D decisions are a result of a political process, we also include instruments that model the political process determining R&D funding. Overidentification tests assume that at least one instrument is valid. Since all energy R&D decisions are made through similar political processes, additional instruments that control for the political pressures that shape R&D decisions help ensure the validity of the overidentification tests. The policy instruments control for a country's general proclivity for government spending and for the political leanings of the government.[11]

V. Results

Our results focus on the two questions posed in the introduction. First, we consider the optimal number of years of lagged energy R&D to include for each technology. With this information in hand, we then examine the impact of both energy R&D and our control variables on energy publications. Next, we link the publications in our data to references on U.S. patents, allowing us to estimate the time it takes for new energy R&D to impact applied research. Finally, we consider the possibility of diminishing returns to large increases in energy R&D, looking both at the quantity of publications (e.g. Does the marginal effect of R&D fall with large increases?) and the quality of publications.

---

[10] Any related technologies are not used as instruments. For example, nuclear R&D is only used as an instrument for biofuels, a transportation-based technology, but not for any technologies pertaining to electricity. Note, for example, that the share of nuclear energy is a control for solar and renewable energy. As such, decisions on nuclear R&D cannot be considered exogenous for these technologies.

[11] For example, Baccini and Urpelainen (2012) show that political shifts affect the volatility of public energy R&D.

A. Estimating the Lag Length

We employ several strategies to identify the appropriate lag length. First, as is common in the literature, the primary criterion is finding the minimum AIC statistic across a range of models (e.g. Crespi and Geuna, 2008).[12] One complication is that our model includes not only lagged values of the R&D variables, but also of several control variables. These controls are often individually insignificant. Thus, we also run models including only a single year of the control and policy variables, to see if this changes the recommended number of lags. We initially examine models including up to ten lags of R&D. Adding additional lags is problematic when the model includes renewable energy policy variables as only two countries in our sample adopted renewable energy certificates before 2003, so that estimates of the lagged value of the REC variable beyond eight years are unreliable. However, adding up to 11 years of lags is possible for biofuels and energy efficiency.[13]

In addition, even collinearity among the R&D variables themselves may cause the AIC statistic to favor smaller lags. Thus, we also run both the full and single control models using a polynomial distributed lag model (PDL). The PDL models provide similar results to the main first differenced models, but by requiring fewer parameters to estimate multiple lags, they offer the potential for lower standard errors on the R&D coefficients. Appendix C presents the AIC statistics and a discussion of their implications for the optimal lag length.

---

[12] Alternatively, we also calculated the BIC statistic for each model, which includes a greater penalty for including irrelevant variables. As a result, collinearity among the lagged R&D values often leads the BIC to recommend fewer lags than the AIC. However, in the case of collinear lagged R&D values, we can still estimate long-run effects that are jointly significant, even when individual year coefficients are estimated imprecisely. Moreover, leaving out relevant lags would lead to omitted variable bias. Thus, I focus on the findings of the AIC statistic in the discussion that follows.

[13] The first US states to adopt REC limits do so in 1998, and Italy adopts an REC limit in 2002. Most states first limits appear in the data in 2003.

In some cases, these strategies recommend different lag lengths. Thus, we also consider the cumulative long-run effect of R&D when evaluating the lag length. These long-run effects, illustrated in Figure 4, tend to be similar across the various estimation techniques, so we focus on the results using first differenced instrumental variables. When the AIC statistic conflicts across models, we use the pattern of cumulative effects as an additional guide. In particular, we consider the lag length at which the cumulative effect of government energy R&D spending levels out, as this is evidence that all relevant lags have been included. Based on these results, we identify an optimal lag length of ten years for biofuels and energy efficiency, six years for solar energy, and seven for wind.

B. Effects of Energy R&D

Having identified the appropriate lag length, Tables 5-6 present results for the first differenced models for each technology. Table 5 shows the results for government R&D for both individual years and the cumulative effect. Table 6 shows the cumulative effects of the various controls included in each model. In each table, the first column for each technology includes results using instruments for current R&D and the second includes results assuming all variables are exogenous. All results include robust standard errors that have been corrected for both heteroskedasticity and autocorrelation.[14]

Before turning to technology-specific results, we first discuss general trends across all models. Public sponsored energy R&D generally leads to increases in publications, although the magnitude of the effect varies across technologies. Other demand characteristics appear less important, as most controls are insignificant, both individually and in the cumulative effects over

[14] Because of the similarities across models, I focus on the first-differenced panel estimates, which avoid imposing any structure on the lag process. Results using the PDL models are included in Appendix D.

time. This contrasts with most research focusing on private sector research efforts, where policy plays an important role (e.g. Johnstone *et al.* 2010, Verdolini and Gaelotti 2011, Dechezleprêtre and Glachant 2014, Nesta *et al.* 2014). Given the long term nature of basic R&D, it is not necessarily surprising that demand factors such as policy shocks have less influence on research supported by the public sector. However, these results also suggest that public R&D funding is not simply replicating support that would otherwise be provided by the private sector. We discuss exceptions to these findings in the technology-by-technology results below.

Turning to the quality of our instruments, the Hansen J test reveals that the instruments are valid in all cases. Each table also presents the p-value of the endogeneity test, where the null hypothesis is that the current value of energy R&D is exogenous. We fail to reject the null for all technologies except solar energy. However, as the results are generally unchanged across endogenous and exogenous specifications, we choose to be conservative and focus on the IV results when presenting the various robustness checks that follow.

Energy R&D has the largest cumulative effect in biofuels. One million dollars of additional government R&D support results in slightly more than two new publications over ten years. One reason for the stronger biofuels effect is the long lag length. While the effect of public R&D levels out for other technologies after six or seven years, for biofuels we find a strong effect even in years nine and ten. There is a strong contemporary effect, with 0.376 publications being induced in year *t*-1. The FD results suggest a cyclical pattern, with strong effects also found in years two and four in the first differenced model before picking up again in year *t*-9. One possibility that the long range effects suggest is that public R&D leads to increases in the supply of scientific personnel and infrastructure devoted to biofuels research. While verifying such an explanation is beyond the scope of the current data set, such questions about the long run impact of public R&D are a fruitful avenue for further study. While most individual controls are insignificant, gasoline prices net of taxes have a large positive impact on publications. Higher prices make substitutes such as biofuels look more attractive and researchers seem sensitive to the market potential of biofuels when deciding on research projects. Also, while the cumulative effect of per capita oil reserves is insignificant, individual year effects have a significant positive effect in years *t*-3, *t*-7, and *t*-8.

Similar to biofuels, the lagged effect for energy efficiency R&D is long. While the largest single year effect occurs in year *t*-2, the magnitudes in year *t*-10 are similar. Compared to the other technologies, the magnitude of both the individual year effects and cumulative effects for energy efficiency are much smaller, with one million of energy efficiency R&D generating just under 0.2 new publications in the long run. Compared to the other technologies in this study, energy efficiency covers a wider range of potential applications, including specific equipment for production, vehicles, and even high technology applications such as computers and server farms. Thus, the smaller magnitude may be an artifact of our data, which relies on keyword searches to identify relevant articles and may miss some of these broader applications. However, it is also possible that, because of the broader nature of energy efficiency research, the costs of generating new energy efficiency results is larger, as the ability to stand on the shoulders of past researchers may be lower in a more diverse research field. Also, while the cumulative effect of per capita oil reserves is insignificant, individual year effects have a significant effect in years *t*-3 through *t*-7, with a positive effect in all years except *t*-5.

Solar energy is the one case in which instrumental variables lead to changes in the results. We reject the null hypothesis that current energy R&D expenditures are exogenous, as the p-value of the endogeneity test is just 0.049. The cumulative effect of energy R&D falls from 2.742 to 2.078 when using instruments for current R&D. The differences are largely driven by changes in the contemporary effect of R&D, which is much larger when exogenous. Recall from the theory section that current variables not only pick up the supply of research projects, but also the demand of editors to publish research on a given topic. As this unobserved demand may be correlated with energy R&D, and that we would expect some time to pass before R&D led to a new publication, it is likely using instrumental variables helps correcting for such demand effects. Policy also plays a role in the case of solar, with larger renewable energy targets inducing more publications.

Finally, for wind energy, the results are consistent across all specifications. In the long-run, a million dollars of energy R&D leads to approximately one new publication. Unlike biofuels, the effect is quick, with most new publications occurring within the first three years. After a brief leveling off, the cumulative effect continues to rise in years six and seven, after which it levels off again, as shown in Figure 4. While none of the controls have a significant cumulative effect, the immediate lag is significant for both the percentage of electricity from hydro and nuclear, although unexpectedly positive for both.

C. Linking basic and applied research – citations by patents

As the ultimate goal of government energy R&D funding is not a publication, but rather a new technology, any evaluation of government R&D should also ask whether any resulting increase in basic science outputs leads to new applications. For this, we link our publication data to patent data, which reflect the output of applied research efforts. Patents contain citations to scientific publications, allowing direct linkages to be made. Moreover, recent work by Roach and Cohen (2013) shows that references to non-patent literature (NPL) such as journal articles are better measures of knowledge flows from public research to patents than are citations to other patents. Using references to non-patent literature found on patents, we gain new information on the speed of diffusion by asking how long it takes energy publications to be cited by patents for new energy technology. The results of this work can thus inform other studies where assumptions about the rate of diffusion of public R&D must be made.

The major challenge for this analysis is linking the publication data used in the previous sections to patents. Due to data constraints, the focus will be on citations made by U.S. patents. Using the International Patent Classification (IPC) system to identify patents pertaining to specific technologies, we identify patents related to biofuels, solar energy, and wind.[15] Data on relevant patents come from the on-line database provided by Delphion (http://www.delphion.com/). We obtained the NPL references for these patents, identifying those referencing journal articles. As there is no standard form for citing articles in a patent, matching articles and patents was done manually.

Table 7 shows the percentage of articles receiving an NPL citation. Because we focus on citations made by U.S. patents, the table separates articles by US and foreign authors, using the same author weights as in the previous section. While the total percentage of articles is about 2 percent, the low number is in because of the large number of articles published in recent years that have yet to be cited. Still, even looking back to articles from earlier five year intervals, the share of articles receiving an NPL citation is always below 10 percent. Moreover, even among articles receiving an NPL citation, the average number of citations received is two or less. Also important is that the opportunity for citation increases over time. Figure 5 shows the total number of patents in each technology by year. For each technology, patenting has increased during the 2000s, as government interest in renewable energy increases (see, e.g. Johnstone *et al*. 2010).

[15] There are no IPC classifications dedicated solely to energy efficiency.

Because of these truncation issues, and because the count of citations is low, we use a hazard regression to focus on the time until the first NPL citation is received. To allow time for articles to be cited, we only consider articles published in 2009 or earlier. Our patent data extends through 2011. The model includes the citation lag, calculated using the publication year of both the cited article and citing patent, a set of country by cited year fixed effects (denoted $\mathbf{YC_{i,t}}$ in the equation below), which control for the different opportunities for future citations available to articles from different countries and from different times, and a dummy variable to control for articles with authors from multiple countries. As we explicitly model the effect of time using the citation lag, we use an exponential baseline hazard:

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

As articles published early in the sample had fewer opportunities for immediate citation, we expect their citation lag to be longer. Thus, in addition to estimating the model using the full sample of articles published between 1991-2009, we also estimate a second model using only articles published from 2000-2009, to ascertain whether the time to citation is faster when citing opportunities increase.

Table 8 presents the regression results, showing that the coefficients on citation lags are significant at the one percent level, except for the squared term on wind energy. To interpret the time to citation, panel A of Figure 6 illustrates how the annual probability of citation changes over time. For the post-2000 sample (shown using solid lines), the annual probability peaks between 8-10 years after publication. The peak shifts out to 11-14 years when considering the full sample (dashed lines). As both the average and median lag between initial application and grant for patents in our sample is 5 years, this means that patents citing these articles are filed 3-5 after publication of the article in the post-2000 sample. Panel B shows the cumulative probability of citation, which begins to grow rapidly 4-6 years after publication, and levels out about 12-14 years afterwards in the post-2000 sample.

While it may take just a few years for an article to be cited by a patent, it also takes a few years for R&D funding to generate new articles. To assess the full time it takes for new energy R&D funding to influence technology development, Figure 7 traces the increased probability of an NPL citation resulting from an additional $1 million R&D funding in year *t*. This calculation combines the regression results from Table 5 to determine the number of new articles induced by additional R&D funding with the results above, predicting the probability of citation to each of these new articles in different years. Appendix E describes the methodology in more detail. Allowing for the lags between initial funding and publication, the probability of a citation resulting from new R&D funding peaks from 10-12 years after funding in the post-2000 sample, and 13-19 years in the full sample. Looking at the cumulative effect, we see little increase in citation until approximately 6 years after funding, with the effect not leveling out until almost 18 years after funding. Again considering a five year window for processing patents, this suggests that new patent applications begin appearing about one year after funding and continue for 13 years. While these lags are shorter than those found by Toole (2012) in his study linking NIH research funding to applications for new molecular entities, as clinical testing prolongs the development of new medicines, these figures still demonstrate that the effect of public energy R&D funding will not be felt until several years after the funding occurs. This finding suggests that papers using just a single value of contemporary or one-year lagged energy R&D to evaluate the effect of public energy R&D spending, such as those cited in section II, do not sufficiently control for the lagged effects between R&D and patenting. While many of these papers do find a positive correlation between current or one-year lagged values of public energy R&D and patenting, the lack of a proper lag structure suggests that these papers are merely picking up endogenous relationships between the factors determining energy R&D funding and those driving renewable energy innovation in the private sector, such as changes in energy prices.

D. Are there diminishing returns to government R&D?

While the previous results provide new information on both the magnitude and duration of the effect of energy R&D, finding a positive effect of public energy R&D funding on related scientific publications is not surprising. However, large increases in energy R&D may come at a cost. Government research funding often fluctuates dramatically. U.S. public energy R&D spending rose from $2.5 billion to $7.8 billion between 1975 and 1982 before leveling of near $3 billion by the mid-1980s. Because researchers have a limited amount of time, and the supply of researchers able and willing to work on a given topic is inelastic in the short run, such large increases may have adjustment costs that limit the potential of large surges in energy R&D funding.

In this section, we consider two additional model specifications to test for the possibility of diminishing returns to research funding: one adding a quadratic term for R&D, and a second including a dummy variable for large increases in R&D. Table 9 presents the cumulative effect for each coefficient, with the model including R&D squared in columns 1-4, and the model using the interaction term in columns 5-8. All models treat current R&D as endogenous[16]. When adding a quadratic term, the marginal effect of R&D will be a function of the R&D spending in a given year. Figure 8 evaluates the marginal effect of energy R&D for a given year for various lags, and Figure 9 shows the cumulative effect. In each, these effects are evaluated for average levels of energy R&D spending, as well as for the 25th, 75th, and 90th percentile. If diminishing returns are

[16] The results are virtually identical if we instead treat current R&D as exogenous.

a concern, we should expect to see lower marginal effects in the upper percentiles.

Our results provide almost no support for diminishing returns. While the quadratic term is negative and significant in a few individual years, the cumulative effect is never negative and significant (columns 1-4 of Table 9). Moreover, it is significant and positive for both biofuels and solar energy raising the possibility of positive spillovers from additional R&D.[17] In Figure 8, we see slightly higher annual marginal effects for the 90th quantile of solar energy R&D for most years, and for biofuels and energy efficiency in the later years. Only for wind do we see the annual marginal effects fall for the highest quantiles, and only in the middle years.

Figure 9 provides more insight, showing how the cumulative effect of energy R&D varies over time for various levels of R&D funding. For biofuels and energy efficiency, the cumulative effect is nearly identical across quantiles until the later years. This would be consistent with the notion that large increases in R&D result in positive spillovers providing dividends in future years, such as by attracting more researchers into the field. For energy efficiency R&D, the cumulative effect for the 90th percentile is below those of other percentiles through year three before catching up, raising the possibility of a period of adjustment for large increases in energy R&D.

Solar energy provides evidence of positive spillovers from large energy increases, with a significant positive quadratic terms. As shown in Figure 9, the cumulative effect of R&D is lowest for the 25th percentile, increasing significantly for each higher percentile. Finally, the only technology showing any evidence of long run diminishing returns is wind. Here, we see that the cumulative effect of wind energy is slightly larger in the 90th percentile in early years, but soon becomes lower than the cumulative effect of other percentiles.

[17] Individual years with significant negative coefficients on the quadratic term include lags 9 and 10 (both positive) for biofuels, lags 2 (positive), 3 (negative), and 8 (positive) for energy efficiency, lags 0, 1, 2, and 5 (all positive) for solar energy, and years 1, 2 (both positive) and 4 (negative) for wind.

Second, we ask specifically whether rapid increases in energy R&D funding are less effective (columns 5-8 of Table 9). We include a dummy variable equal to one if energy R&D increases by 100% or more in a given year, and interact this variable with the level of energy R&D funding. Thus, a negative coefficient on this interaction indicates that the marginal effect of R&D is lower in years with a doubling or more of R&D. As Figure 10 shows, the number of countries choosing to double energy R&D in a given year is generally equally dispersed across time, with one or two countries per year experiencing such increases. The one exception is the large number of countries choosing to double energy R&D as part of stimulus packages in 2009 and 2010. Overall, between 6 and 12 percent of all country/year observations from 1981-2011 include a doubling of energy R&D.

The results confirm that diminishing returns and adjustment costs are not a significant concern given the current levels of energy R&D funding and provide some support for positive spillovers. The coefficient on the interaction term is positive for all technologies except solar, with statistically significant results for biofuels and wind. For individual years, the interaction is positive and significant at the 5% level for several years in the biofuels and wind regressions. Thus, despite concerns about diminishing returns to R&D, we instead find evidence that a doubling of energy R&D creates potential positive spillovers that makes energy R&D spending more productive.

E. Does Quality Change: Citation Analysis

While we find little evidence of diminishing returns regarding the number of publications generated from increased government R&D, these results tell us nothing about the quality of those publications. Even if publication counts increase as funding increases, those projects only able to win support during periods of ample funding would be expected to have lower quality than those projects earning support even when available funding is low. Thus, diminishing returns to research may be exhibited not in the quantity of publications, but in the quality of publications. To assess publication quality, we turn to citation data. The assumption made is that more frequently cited articles are of higher quality. Using citations as an indicator of article quality is a common technique in bibliometric analyses (see, for example, National Science Board, 2008). Within economics, patent citations have been used in a similar way (e.g. Trajtenberg 1990, Lanjouw and Schankerman 2004, Popp 2002, 2006).

As in the patent citation literature, simple counts of citations received are not sufficient to evaluate the quality of an article. The number of citations that an article receives depends not only on the quality of the article, but also on the number of opportunities for citation. Let *i* represent an article, *j* represent the home country of the article's authors, and *t* represent the publication year of the article. The number of citations received by an article from country *j* and published in year *t* can be described by the following relationship:

$$(2) \qquad NumCite_{i,j,t} = f(\ln(NUMPubs_{j,t}),\ \ln(NUMPubs_{j,t})^2,\ \alpha_j, \beta_t).$$

α_j and β_t represent country and year fixed effects. We expect articles to receive more citations from other authors in the same country. Country fixed effects acknowledge that the number of citing opportunities may vary by country. Year fixed effects capture the number of citing opportunities that occur in the years after publication. For energy publications, we would expect the raw number of citations to be higher during periods of intense research activity, such as when energy prices are highest.

Our variable of interest is $NUMPubs_{j,t}$. This represents the total number of articles published on a given topic in country *j* in a given year, *t*. If additional research leads to marginal articles being published, we would expect the number of citations to fall when more articles are published. To allow for the possibility of positive spillovers at moderate levels of research, we also estimate a model using a squared term. If positive spillovers exist at moderate levels of research, we would expect the linear term to be positive, with the squared term being negative. Because the average number of publications per year varies by country, we use the log of publications in the regressions that follow. Because many articles receive zero citations, we use a generalized negative binomial regression to estimate equation (4). All standard errors are clustered by article.

One final complication is that just over 25% of all articles include authors from multiple countries. For these cases, we include a separate observation for each article/country pair, using weighted regression to weight each article/country pair by the share of authors from that country. For example, if an article has two authors from the United States and one from Canada, we would include observations for both the United States and Canada, with a weight of 0.67 for the U.S. observations and a weight of 0.33 for the Canadian observations. We include a separate dummy variable for these articles, as we expect them to receive more citations than single-country authors, since the authors of these articles are exposed to multiple research networks.[18]

Table 10 presents the results. Publication counts are in logs, so that the coefficients can be interpreted as elasticities. For all technologies except solar energy, we find evidence of diminishing returns, with an elasticity between publications and citations ranges from -0.167 to -0.186. Results are significant at the 5% level for each of these three technologies. Including a squared term provides no evidence of increasing returns for smaller numbers of publication. As

[18] As a robustness check, we also run our regressions dropping all multicountry articles. General trends are the same in both models, although the results for wind are no longer significant when dropping multicountry articles. These results are available upon request.

expected, multicountry articles are 17-24% more likely to be cited than articles with authors from a single country.

VI. Conclusion

Although government R&D support for alternative energy and energy efficiency improvements is seen as an important component of climate change policy, empirical evidence on the effectiveness of such R&D is limited. Attempts to assess the effectiveness of government energy R&D support are complicated by the long lags between initial funding and the final research outcomes, as well as by the challenge of isolating the effect of public R&D funding from demand-side influences, such as energy prices and environmental policy, that influence both public and private R&D activity. In this paper, we use scientific publications as an indicator of basic research output to assess the effectiveness of energy R&D. Using a panel of OECD countries, we control for other factors that may influence the direction of research and use instrumental variables to separately identify the effect of publicly funded R&D on publications.

The results show that, even controlling for other energy policies and energy prices that will influence private R&D funding decisions, government R&D support does increase the number of related energy publications. In general, an additional million dollars of energy R&D leads to 1-2 additional publications. Moreover, factors found to influence private R&D activity in other papers, such as energy prices and policy, have little impact on publications. Thus, it does not appear that public R&D merely substitutes for other sources of funding.

This effect of energy R&D occurs over a period of years, with lags as long as 10 years in the case of biofuels and energy efficiency. Moreover, as the ultimate goal of public R&D is to provide the building blocks for new innovation, we use citations to link the publication data to

U.S. patents, providing evidence of the time it takes for public R&D to stimulate private sector innovation. Related patent applications begin soon after funding and continue for up to 13 years. That long lags matter is important to the evaluation of energy R&D, as many studies evaluating the effect of government R&D on innovation consider just contemporary or one-year lagged energy R&D. While these studies often find little effect of energy R&D on private sector innovation, failure to consider sufficient lags call these results into question.

Importantly, we also find little evidence of diminishing returns to public energy R&D. We find no evidence that the marginal increase in publications from new R&D funding declines following large increases in energy R&D. We do find some evidence of diminishing quality, as citations to new articles fall by approximately 1.7% when energy R&D increases by 10 percent for three of the four technologies studied. Thus, while the number of publications induced continues to increase as energy R&D expands, large increases in funding may simply provide support for marginal projects of less overall value to society. Overall, the lack of evidence for diminishing returns to energy funding suggests that there is room to expand current public energy R&D efforts, and that the constraints for funding are likely to come from other sources, such as macroeconomic constraints or the pool of scientist and engineering personnel currently available to work on energy projects. In particular, the long-run effects particularly evident in the case of biofuels and energy efficiency suggest that increases in public funding have positive spillovers, either through increasing the amount of research on which to build or by increasing the number of researchers in the field. Additional research to better understand the mechanisms behind these long run effects is warranted.

References

Adams, James D. (1990), “Fundamental Stocks of Knowledge and Productivity Growth,” *Journal of Political Economy*, 98(4), 673-702.

Adams, James D. and J. Roger Clemmons (2008), “Science and Industry: Tracing the Flow of Basic Research through Manufacturing and Trade,” *Economics of Innovation and New Technology*,” 17(5), 473 - 495.

Adams, James D., J. Roger Clemmons, and Paula E. Stephan. (2004), “Standing on Academic Shoulders: Measuring Scientific Influence in Universities,” *NBER Working Paper #10875*.

Adams, James D., Clemmons, J. Roger, and Stephan, Paula E. (2006), “How rapidly does Science leak out?” *NBER Working Paper #11997*.

Almon, Shirley. 1965. The distributed lag between capital appropriations and expenditures. *Econometrica* 33: 178-196

Azoulay, Pierre, Joshua Graff Zivin, and Jialan Wang. 2010. Superstar extinction. *Quarterly Journal of Economics* 125(2): 549-589.

Baccini, Leonardo and Johannes Urpelainen. 2012. Legislative fractionalization and partisan shifts to the left increase the volatility of public energy R&D expenditures. *Energy Policy* 46: 49-57.

Bhattacharya, Jay and Mikko Packalen. 2011. Opportunities and benefits as determinants of the direction of scientific research. *Journal of Health Economics* 30: 603-615.

Blume-Kohout, Margaret E. 2012. Does targeted, disease-specific public research funding influence pharmaceutical innovation? *Journal of Policy Analysis and Management* 31(3), 641-660.

Cameron, A. Colin and Pravin K. Trivedi. 2009. *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.

Crespi, Gustavo A. and Aldo Geuna (2008), "An empirical study of scientific production: A cross country analysis, 1981-2002," *Research Policy*, 37, 565-579.

Dechezleprêtre, A, Popp, D. 2015. Fiscal and regulatory instruments for clean technology development in the European Union. *CESifo Working Paper # 5361*.

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

Ek, Kristina and Patrik Söderholm. 2010. Technology learning in the presence of public R&D: The case of European wind power. *Ecological Economics*, 69, 2356-2362.

Freeman Richard and John Van Reenen (2009), "What if Congress doubled R&D spending on the physical sciences?" in Josh Lerner, Scott Stern, eds. *Innovation Policy and the Economy, vol. 9*, University of Chicago Press, Chicago, IL, 1-38.

Furman, Jeffrey L., Fiona Murray, and Scott Stern. 2012. Growing stem cells: The impact of federal funding policy on the U.S. scientific frontier. *Journal of Policy Analysis and Management* 31(3), 661-705.

Goolsbee, Austan. 1998. Does government R&D policy mainly benefit scientists and engineers? *American Economic Review* 88(2): 298-302.

Im, K.S., M.H. Pesaran, and Y. Shin. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115: 53-74.

International Energy Agency. 2014. *CO_2 Emissions From Fuel Combustion: Highlights*. OECD/IEA, Paris, France.

International Energy Agency. 2011. *IEA Guide to Reporting Energy RD&D Budget/Expenditure Statistics*. OECD/IEA, Paris, France.

Jacob, Brian A, Lars Lefgren. 2011. The impact of research grant funding on scientific productivity. *Journal of Public Economics* 95: 1168-1177.

Johnstone, Nick, Ivan Hascic, and David Popp (2010), "Renewable energy policies and technological innovation: Evidence based on patent counts," *Environmental and Resource Economics*, 45(1), 133-155

Jones, Benjamin F., Stefan Wuchty, and Brian Uzzi (2008), "Multi-university research teams: Shifting impact, geography, and stratification in science," *Science*, 322, 1259-1262.

Klaassen, G., S. Miketa, K. Larsen, and T. Sundqvist. 2005. The impact of R&D on innovation for wind energy in Denmark, Germany and the United Kingdom, *Ecological Econ.* **54**, 227-240.

Lane, Julia (2009), "Assessing the impact of science funding," *Science*, 324, 1273-1275.

Lanjouw, Jean O. and Mark Shankerman (2004), "The Quality of Ideas: Measuring Innovation with Multiple Indicators," *Economic Journal*, 114(495), 441-465.

National Science Board (2008), *Science and Engineering Indicators 2008*, Arlington, VA: National Science Foundation (NSB 08-01).

Nesta, Lionel, Francesco Vona, and Francesco Nicolli (2014), "Environmental policies, competition, and innovation in renewable energy," *Journal of Environmental Economics and Management*, 67, 396-411.

OECD-EPAU (2013) Renewable Energy Policy Dataset, version February 2013. Compiled by the OECD Environment Directorate's Empirical Policy Analysis Unit (Johnstone, N., Haščič, I., Cárdenas Rodríguez M., Duclert, T.) in collaboration with an ad hoc research consortium (Arnaud de la Tour, Gireesh Shrimali, Morgan Hervé-Mignucci, Thilo Grau, Emerson

Reiter, Wenjuan Dong, Inês Azevedo, Nathaniel Horner, Joëlle Noailly, Roger Smeets, Kiran Sahdev, Sven Witthöft, Yunyeong Yang, Timon Dubbeling).

Pardy, Philip G. (1989), "The agriculture knowledge production function: An empirical look," *Review of Economics and Statistics*, 71(3), 453-461.

Popp, David, Ivan Hascic and Neelakshi Medhi (2011), "Technology and the Diffusion of Renewable Energy," *Energy Economics*,, 33(4), 648-662

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

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

Popp, David (2001), "The Effect of New Technology on Energy Consumption," *Resource and Energy Economics*, 23(3), 215-239.

Roach, Michael and Wesley M. Cohen, (2013). Lens or Prism? Patent Citations as a Measure of Knowledge Flows from Public Research. *Management Science* 59(2):504-525. http://dx.doi.org/10.1287/mnsc.1120.1644

Rosenbloom, Joshua L., Donna K. Ginther, Ted Juhl, and Joseph Heppert (2014), "The effects of research and development funding on scientific productivity: Academic chemistry 1990-2009," *NBER Working Paper #20595*.

Schuelke-Leech, Beth-Anne. 2014. Volatility in federal funding of energy R&D. *Energy Policy* 67, 943-950.

Söderholm Patrik and Ger Klaassen (2007), "Wind Power in Europe: A Simultaneous Innovation-Diffusion Model," *Environmental and Resource Economics*, 36, 163-190.

Söderholm Patrik and T. Sundqvist (2007), "Empirical challenges in the use of learning curves for assessing the economic prospects of renewable energy technologies," *Renewable Energy*. 32, 2559-78.

Thorsten Beck, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh, 2001. New tools in comparative political economy: The Database of Political Institutions. 15:1, 165-176 (September), *World Bank Economic Review*.

Toole, Andrew A. 2012. The impact of public basic research on industrial innovation: Evidence from the pharmaceutical industry. *Research Policy* 41, 1-12.

Trajtenberg, Manuel (1990), "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *RAND Journal of Economics*, 20, 172-187.

Verdolini, Elena and Marzio Galeotti. 2011. At home and abroad: An empirical analysis of innovation and diffusion in energy technologies. *Journal of Environmental Economics and Management* 61, 119-134.

Wooldridge, Jeffrey M. 2012. *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, MA.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi (2007), "The increasing dominance of teams in production of knowledge," *Science*, 316, 1036-1039.

Zucker, Lynne G., Michael R. Darby, Jonathan Furner, Robert C. Liu, and Hongyan Ma (2007), "Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production," *Research Policy*, 36(6), 850-863.

Figure 1 – Trends in Energy R&D and Publications – All Countries

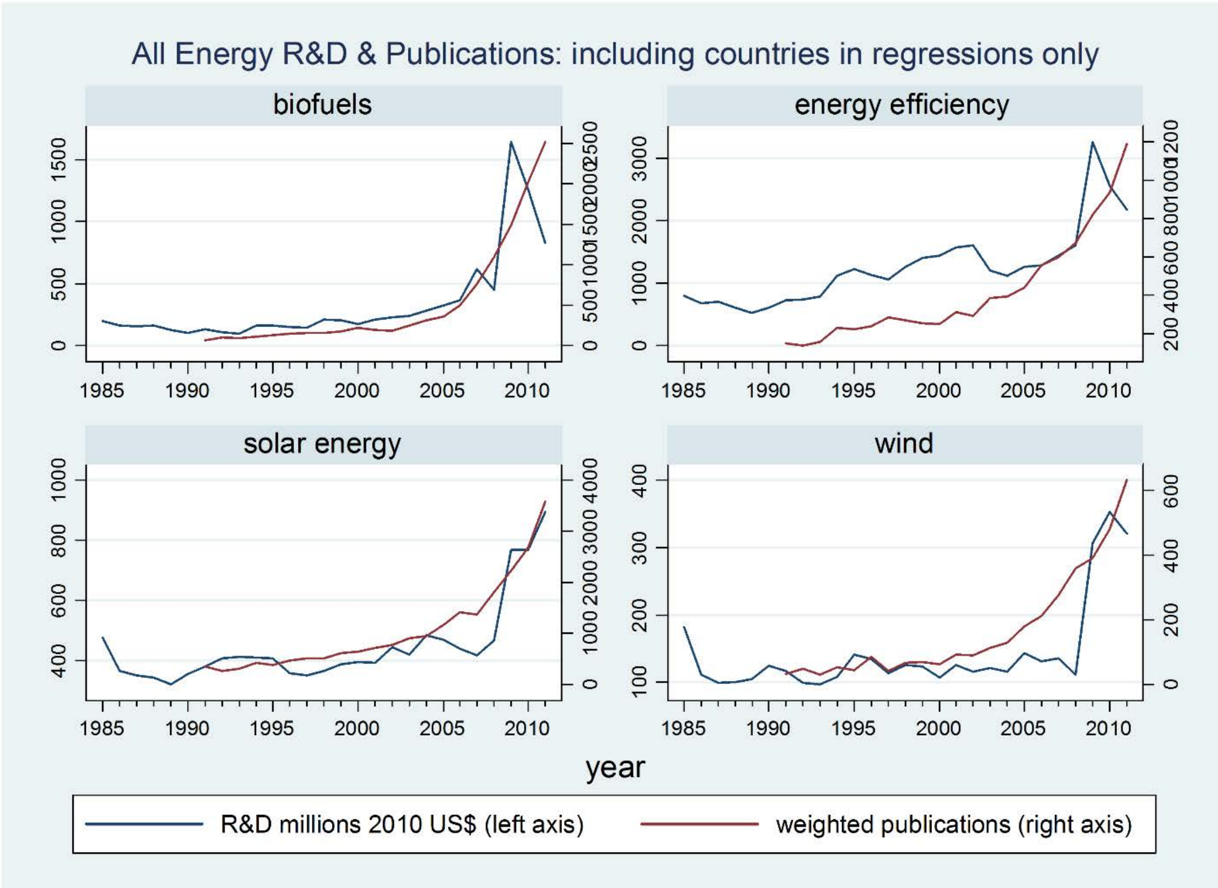


Figure 2 – Trends in Energy R&D and Publications by Country: Solar Energy

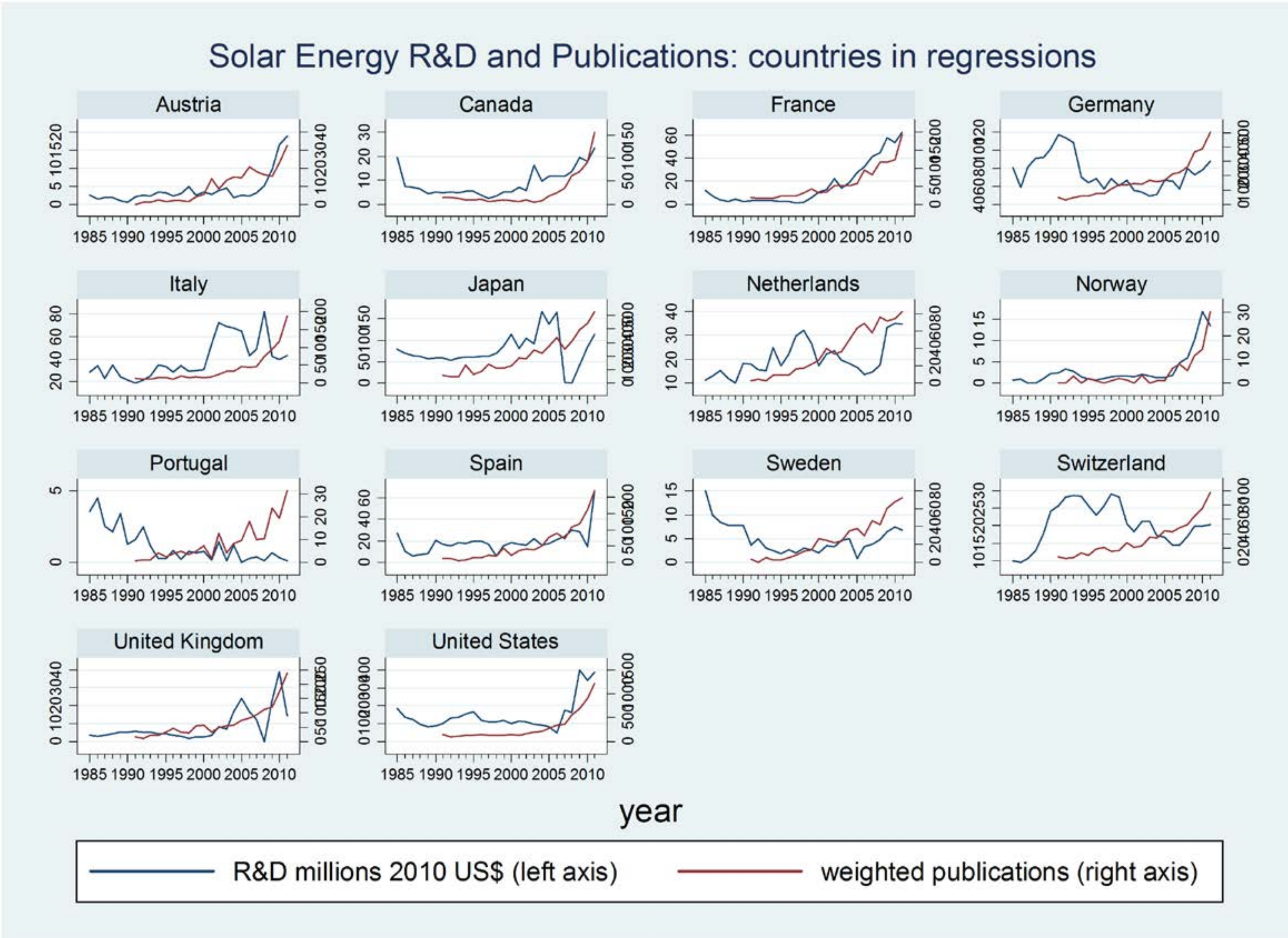


Figure 3 – Trends in Energy R&D and Publications by Country: Wind Energy

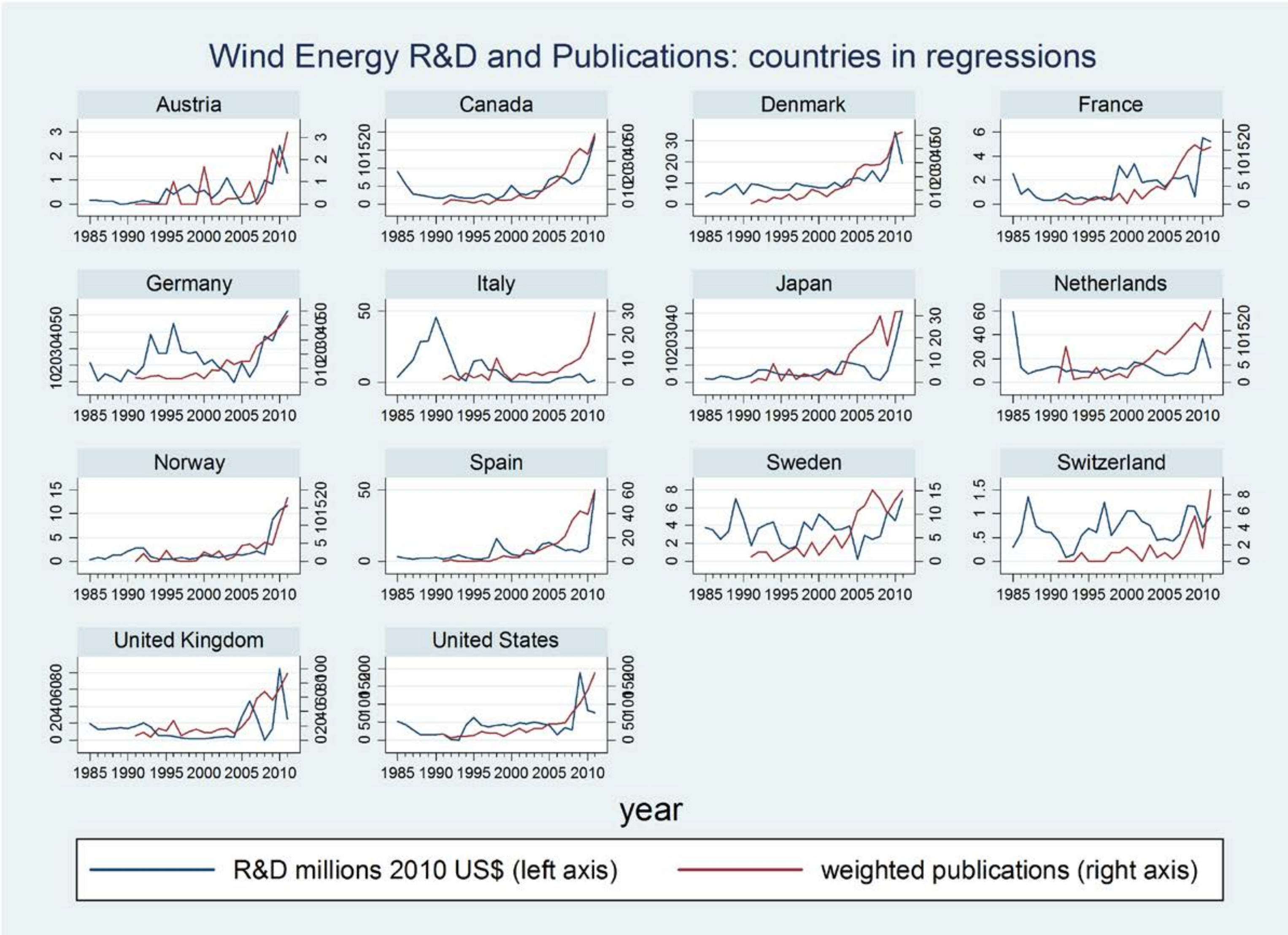


Figure 4 – Cumulative effect of energy R&D on publications

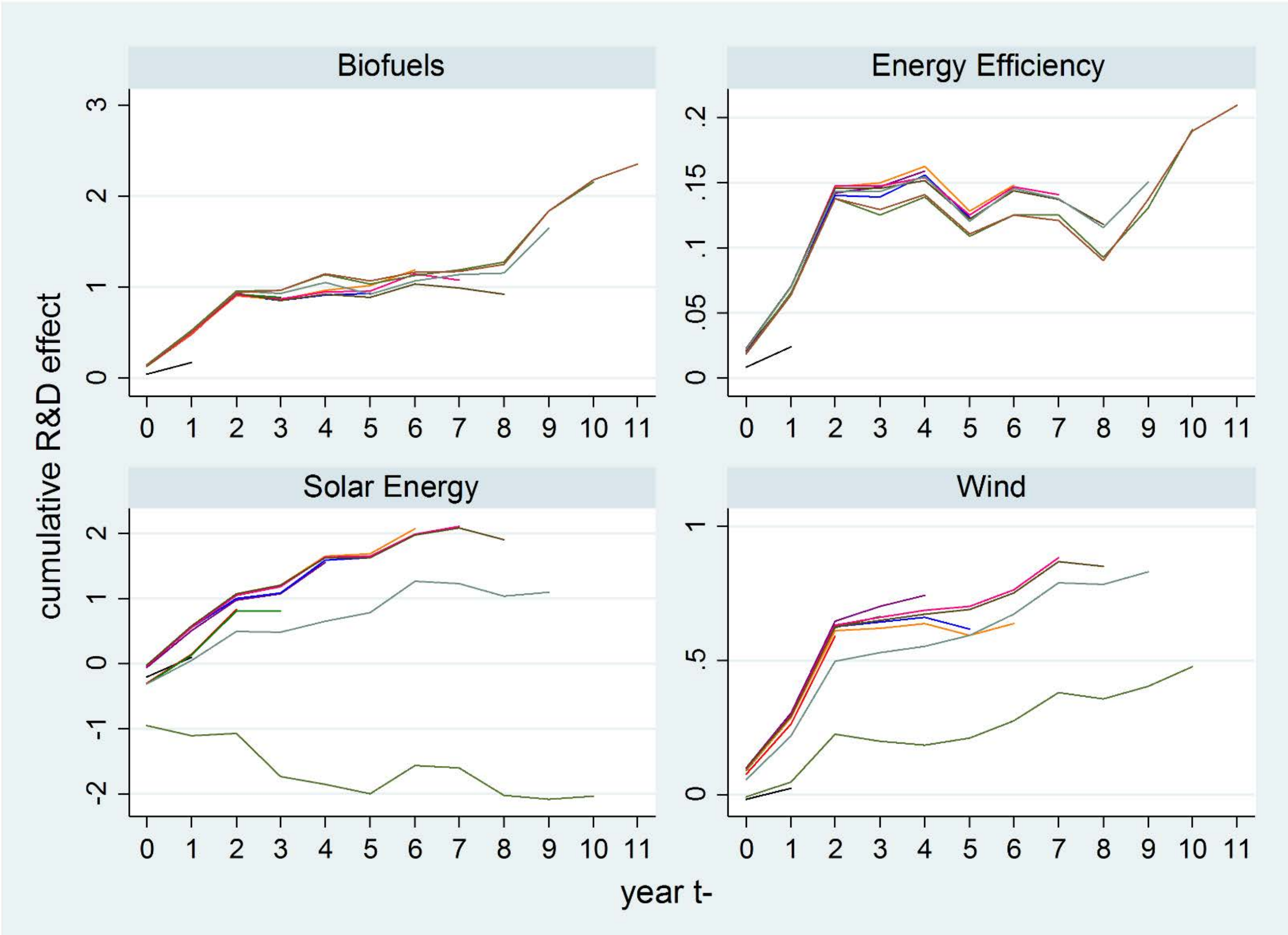


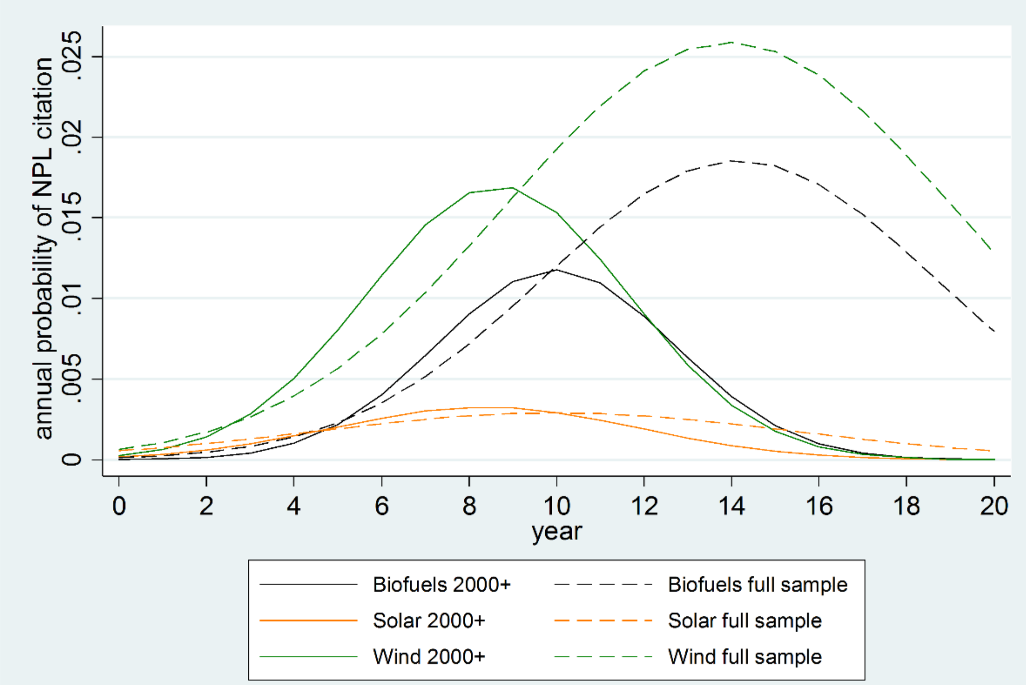
Figure shows the cumulative effect of energy R&D on publications through year t-x, where x is the year shown on the x-axis. Calculated using the first differenced model with instrumental variables for contemporary energy R&D.

Figure 5 – Alternative energy patents over time

Figure shows the total number of patents granted in the U.S. for biofuels, solar, and wind energy. Patents are dated using their priority year.

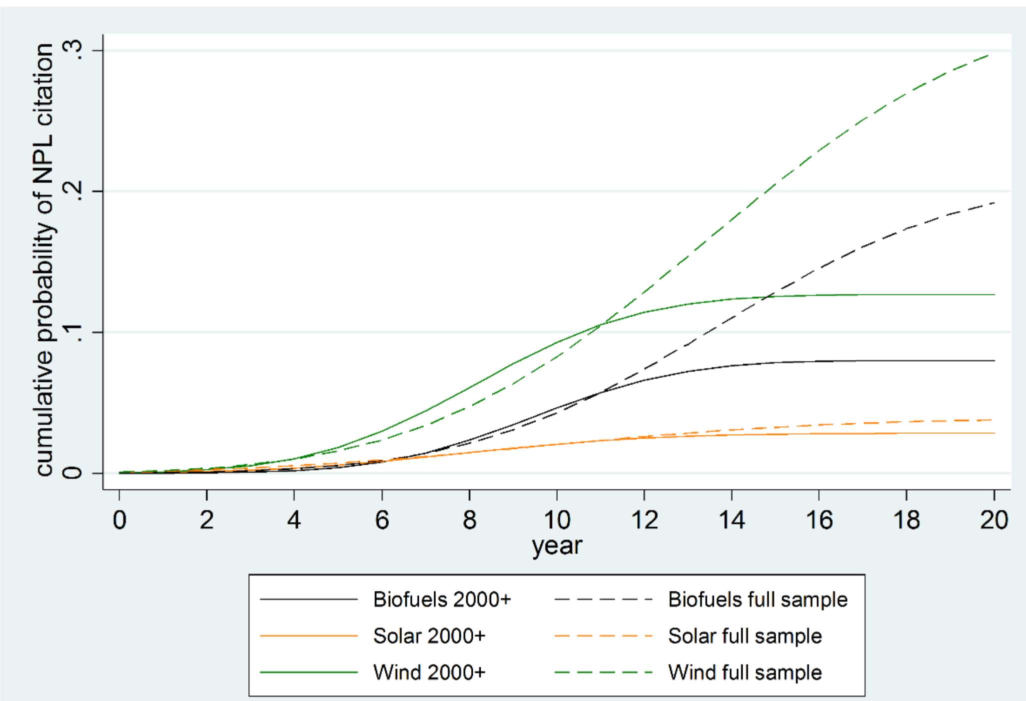
Figure 6 – Probability of NPL citation over time

A. Annual probability of NPL citation



NOTE: Figure shows annual probability of NPL citation, by a single article t years after publication

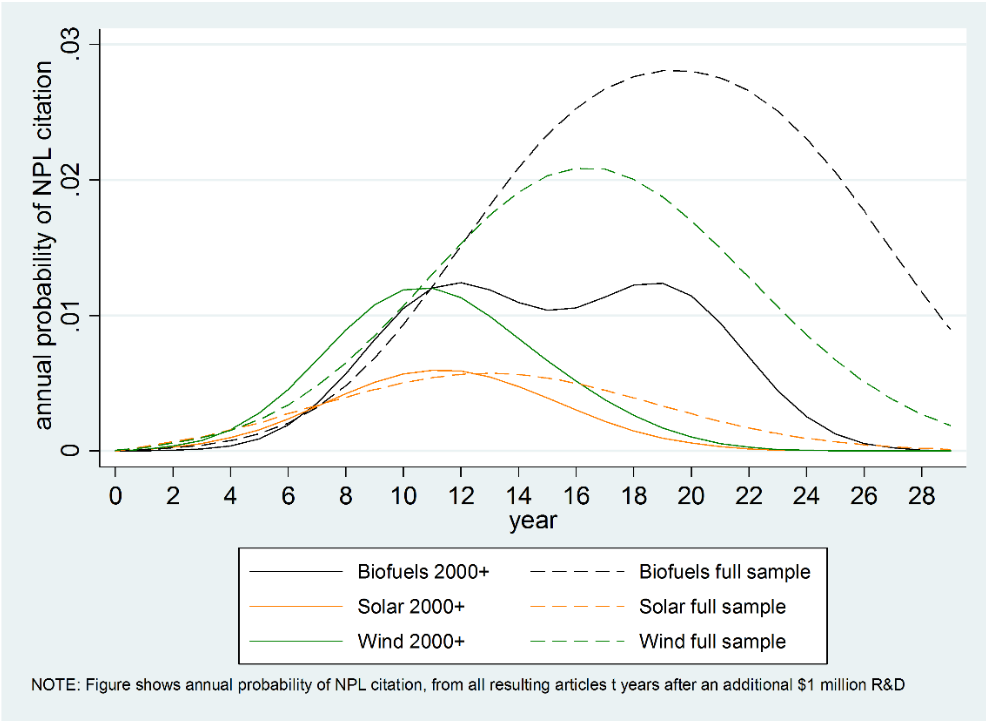
B. Cumulative probability of NPL citation



NOTE: Figure shows cumulative probability of NPL citation, by a single article t years after publication

Figure 7 – Increased probability of NPL citation from additional energy R&D

A. Annual probability of NPL citation



B. Cumulative probability of NPL citation

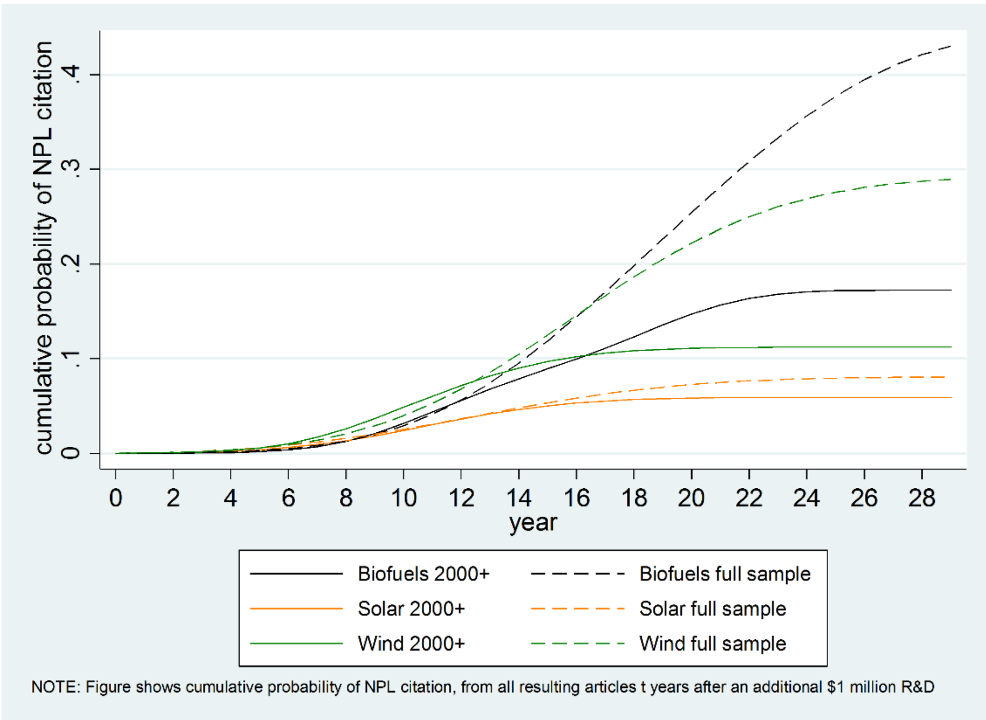


Figure 8 – Marginal effect of R&D per year: Models including R&D squared

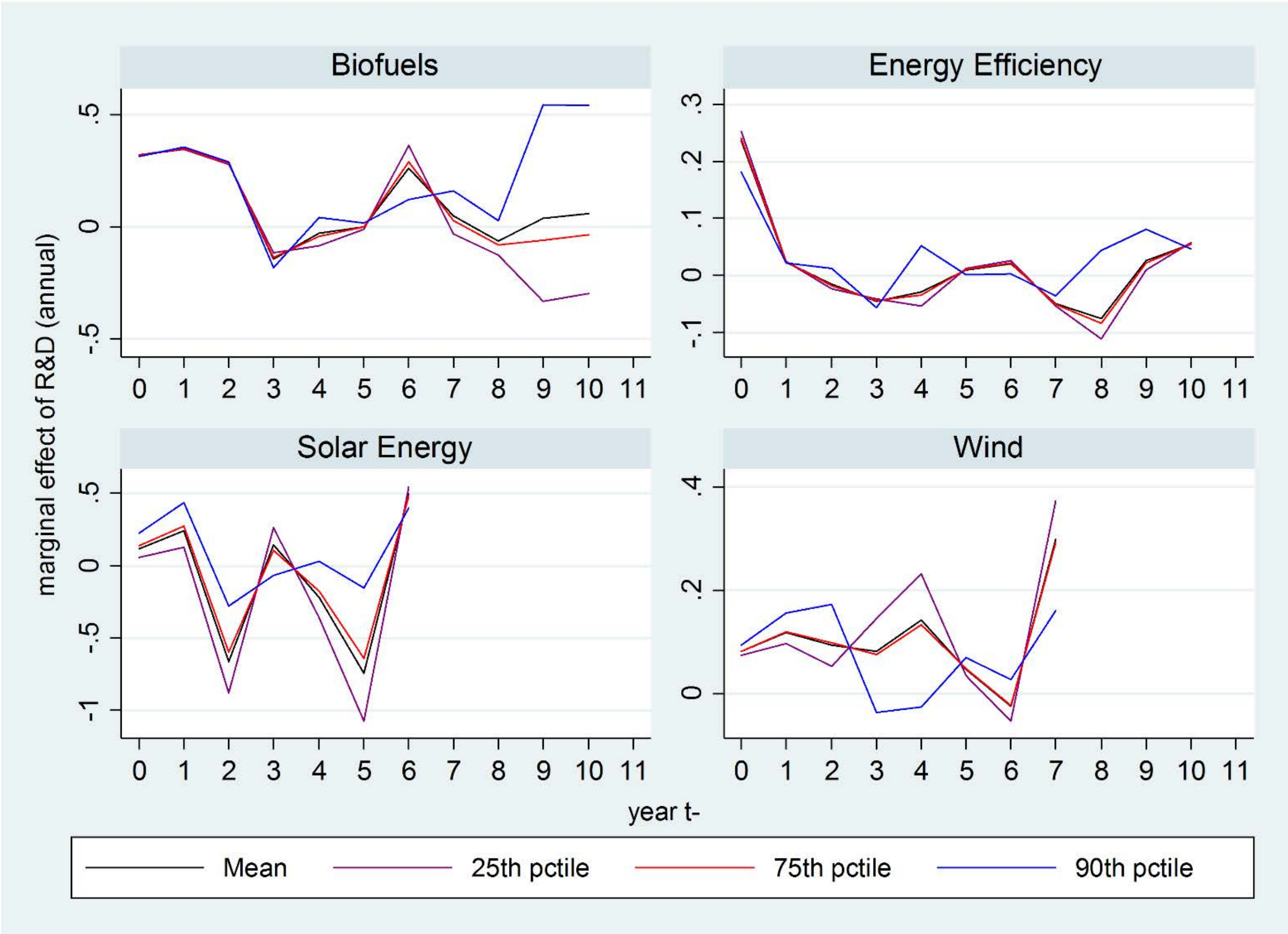


Figure shows the marginal effect of energy R&D on publications in year t-x, where x is the year shown on the x-axis. Calculated at different percentiles of energy R&D spending, using the first differenced model including R&D^2 with instrumental variables for contemporary energy R&D.

Figure 9 – Cumulative effect of R&D: Models including R&D squared

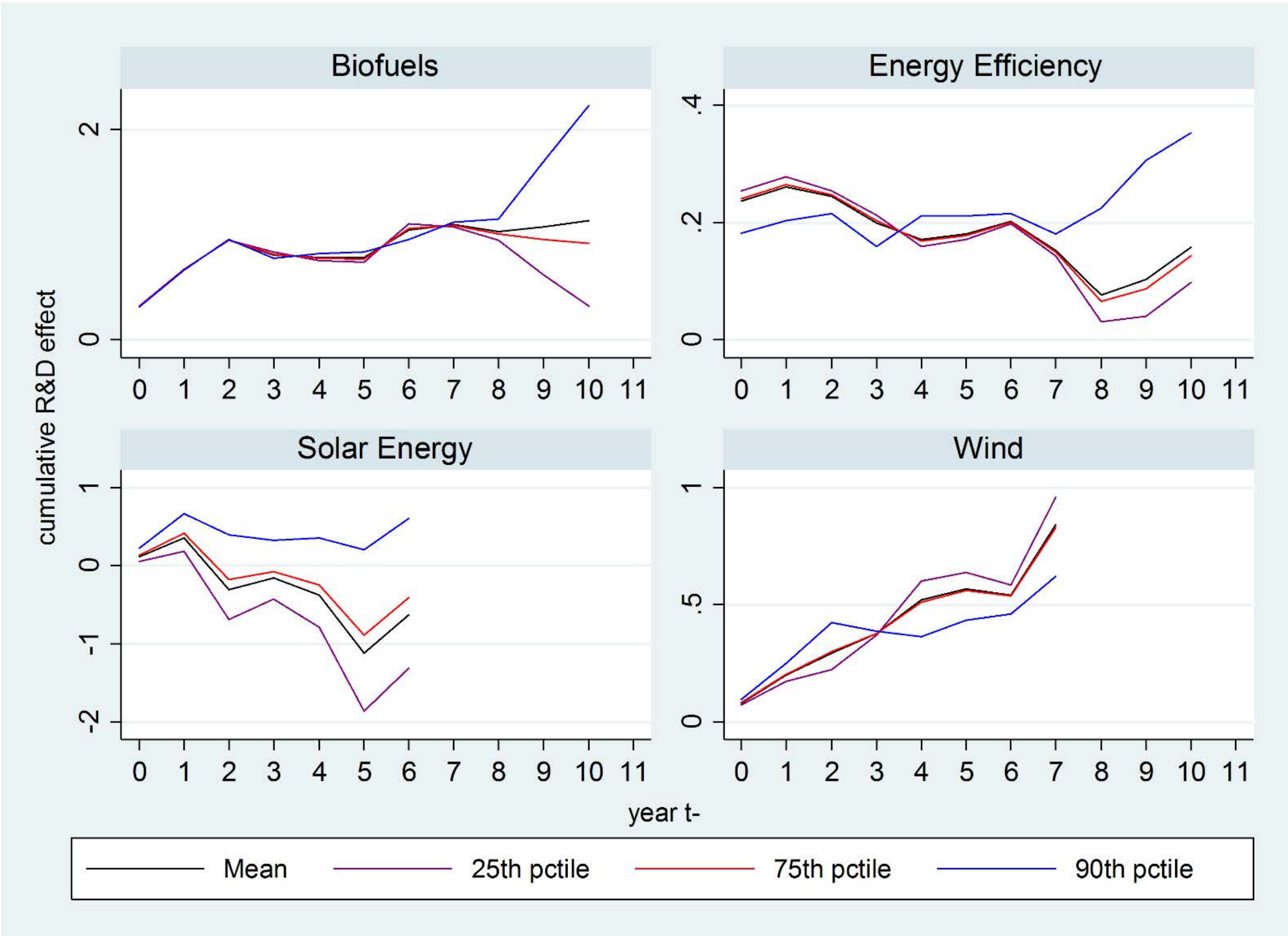


Figure shows the cumulative marginal effect of energy R&D on publications through year t-x, where x is the year shown on the x-axis. Calculated at different percentiles of energy R&D spending, using the first differenced model including R&D^2 with instrumental variables for contemporary energy R&D.

Figure 10—Number of countries doubling energy R&D per year

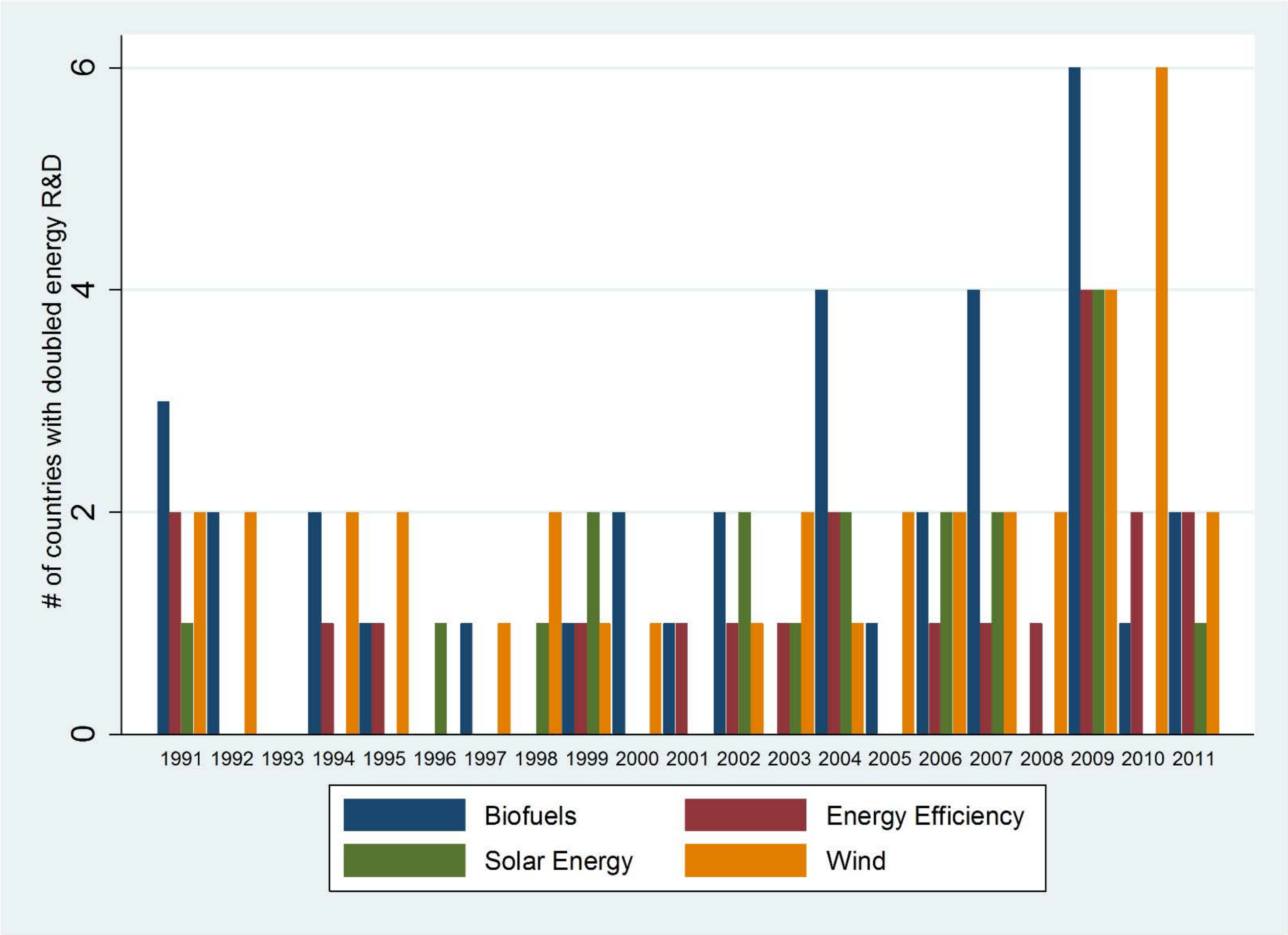


Figure shows the number of countries whose energy R&D budget for a given technology is at least double that of the previous year.

Table 1 – Countries included for each technology

	Biofuels	Energy Efficiency	Solar Energy	Wind
Austria	X	X	X	X
Canada	X	X	X	X
Denmark	X	X		X
France	X	X	X	X
Germany	X	X	X	X
Italy	X	X	X	X
Japan	X	X	X	X
Netherlands	X	X	X	X
New Zealand		X		
Norway	X	X	X	X
Portugal	X	X	X	
Spain	X	X	X	X
Sweden	X		X	X
Switzerland	X	X	X	X
United Kingdom	X		X	X
United States	X	X	X	X
TOTAL	15	14	14	14

Table 2 – Descriptive Statistics for R&D and Publications

technology	variable	N	mean	sd	min	max
biofuels	R&D	315	25.48	86.02	0.00	1186.90
	weighted publications	315	35.06	96.38	0.00	1138.25
energy efficiency	R&D	294	102.31	216.09	0.01	2149.17
	weighted publications	294	29.94	55.95	0.00	493.99
solar energy	R&D	294	33.49	51.17	0.00	401.70
	weighted publications	294	74.66	128.47	0.00	1208.84
wind	R&D	294	10.69	17.54	0.00	187.19
	weighted publications	294	11.83	20.03	0.00	188.20
Total	R&D	1197	42.69	123.79	0.00	2149.17
	weighted publications	1197	37.82	88.68	0.00	1208.84

NOTES: R&D in millions of 2010 US $

Table 3 – Top 10 publication sources, by technology

Biofuels	
United States	4428.62
Peoples R China	1817.09
India	1279.08
Brazil	949.75
Turkey	793.54
Japan	761.61
United Kingdom	749.06
Canada	735.17
Germany	735.05
Spain	714.40

Energy Efficiency	
United States	3753.56
Peoples R China	1932.32
Japan	1308.66
South Korea	839.05
United Kingdom	834.05
Germany	719.78
Canada	634.77
Italy	613.12
Taiwan	545.87
France	477.72

Solar Energy	
United States	6323.95
Peoples R China	5044.19
Japan	4314.37
Germany	3525.28
South Korea	2415.65
India	2123.81
Taiwan	1506.67
United Kingdom	1409.72
France	1219.91
Spain	1187.10

Wind	
United States	916.47
United Kingdom	571.13
Denmark	337.15
Germany	290.77
Spain	268.43
Peoples R China	255.60
Canada	251.87
Japan	222.81
Greece	200.63
Turkey	197.62

Table 4 – Other variables and data sources

Variable	Source	Biofuels	Energy Efficiency	Solar	Wind
Policy Variables					
Gasoline taxes (2010 US $ per liter)	IEA	X			
% renewables required by RPS	OECD-EPAU			X	X
Feed-in tariff, solar PV (2010 US $ per kWh)	OECD-EPAU			X	
Feed-in tariff, wind (2010 US $ per kWh)	OECD-EPAU				X
Control Variables					
ln(per capita GDP) (2005 US$)	OECD	X	X	X	X
Gasoline prices, w/out taxes (2010 US $ per liter)	IEA	X			
Gasoline prices, inc. taxes (2010 US $ per liter)	IEA		X		
Household electricity price (2010 US$ per MWh)	IEA		X		
Crude oil proved reserves per capita (million barrels)	EIA/OECD[a]	X	X		
Natural gas proved reserves per capita (billion cu. ft.)	EIA/OECD[a]		X		
Coal production per capita (short tons)	EIA/OECD[a]		X		
% electricity from hydropower	EIA			X	X
% electricity from nuclear	EIA			X	X
growth rate of electricity consumption	EIA			X	X
Instruments					
Dummy: is executive branch party orientation right-wing	DPI	X	X	X	X
Dummy: is executive branch party orientation center	DPI	X	X	X	X
Tax revenue (excluding social security), % of GDP	OECD	X	X	X	X
General government consumption expenditure, % of GDP	OECD	X	X	X	X
Government energy R&D: biofuels	IEA			X	X
Government energy R&D: energy efficiency	IEA	X			X
Government energy R&D: energy storage	IEA	X	X		
Government energy R&D: nuclear fusion	IEA	X			
Government energy R&D: solar energy	IEA		X		
Government energy R&D: wind	IEA	X	X		

DPI: Database of Political Institutions 2012 (Beck et al 2001); EIA: US Energy Information Administration; IEA: International Energy Agency Energy Prices and Taxes Database (IEA 2006); OECD: OECDStat; OECD-EPAU: OECD Renewable Energy Policy Database (OECD-EPAU 2013) , an update and extension of the dataset originally used in Johnstone et al. (2010).

[a]: Reserve data from Energy Information Administration; population data to calculate per capita values from OECDStat

Table 5 – First-differenced panel regression results: Government R&D

	Biofuels		Energy Efficiency		Solar Energy		Wind Energy	
	IV	exog	IV	exog	IV	exog	IV	exog
RD	0.152***	0.152***	0.0197***	0.0113**	-0.0267	0.372***	0.0983**	0.124***
	(0.0153)	(0.0151)	(0.00649)	(0.00524)	(0.276)	(0.0839)	(0.0440)	(0.0246)
RD(t-1)	0.376***	0.376***	0.0460***	0.0443***	0.601***	0.663***	0.197***	0.212***
	(0.0259)	(0.0245)	(0.00675)	(0.00705)	(0.167)	(0.150)	(0.0520)	(0.0462)
RD(t-2)	0.429***	0.429***	0.0722***	0.0697***	0.483**	0.331*	0.335***	0.348***
	(0.0196)	(0.0196)	(0.00886)	(0.00845)	(0.246)	(0.200)	(0.0547)	(0.0549)
RD(t-3)	0.00719	0.00721	-0.0128	-0.0185	0.154	0.312**	0.0313	0.0525
	(0.0719)	(0.0720)	(0.0296)	(0.0294)	(0.144)	(0.127)	(0.0580)	(0.0587)
RD(t-4)	0.176**	0.176**	0.0141	0.0190	0.447**	0.601***	0.0263	0.0287
	(0.0891)	(0.0891)	(0.0253)	(0.0253)	(0.191)	(0.225)	(0.0585)	(0.0575)
RD(t-5)	-0.101	-0.100	-0.0303	-0.0324	0.0350	0.116	0.0130	0.0216
	(0.109)	(0.109)	(0.0255)	(0.0253)	(0.243)	(0.225)	(0.0532)	(0.0523)
RD(t-6)	0.0940	0.0942	0.0167	0.00891	0.385**	0.347*	0.0618	0.0639
	(0.104)	(0.104)	(0.0176)	(0.0165)	(0.194)	(0.186)	(0.0519)	(0.0515)
RD(t-7)	0.0585	0.0588	0.000165	0.00361			0.121***	0.127***
	(0.0965)	(0.0956)	(0.0229)	(0.0222)			(0.0451)	(0.0436)
RD(t-8)	0.0884	0.0882	-0.0329	-0.0331				
	(0.118)	(0.116)	(0.0255)	(0.0249)				
RD(t-9)	0.558***	0.558***	0.0381	0.0408				
	(0.160)	(0.162)	(0.0268)	(0.0266)				
RD(t-10)	0.318**	0.319**	0.0598***	0.0616***				
	(0.152)	(0.151)	(0.0211)	(0.0212)				
Cumulative effects:								
R&D	2.155***	2.158***	0.191***	0.175***	2.078***	2.742***	0.883***	0.977***
	(0.348)	(0.353)	(0.0520)	(0.0511)	(0.599)	(0.573)	(0.225)	(0.202)
N	300	300	280	280	280	280	280	280
AIC	2351.3	2351.3	2110.7	2108.3	2676.4	2648.1	1766.5	1765.3
BIC	2703.2	2703.2	2535.9	2533.5	2999.9	2971.5	2115.5	2114.3
F RD 1st stage	32.51		18.01		11.40		38.49	
Hansen J p-value	0.293		0.183		0.491		0.502	
Endog. test p-value	0.979		0.240		0.0459		0.258	

Standard errors in parentheses. All models use robust standard errors with correction for autocorrelation.
*: significant at 10% level. **: significant at 5% level. ***: Significant at 1% level.

Table 6 – First differenced panel regression results: Controls

	Biofuels		Energy Efficiency		Solar Energy		Wind Energy	
	IV	exog	IV	exog	IV	exog	IV	exog
Cumulative effects:								
R&D	2.155***	2.158***	0.191***	0.175***	2.078***	2.742***	0.883***	0.977***
	(0.348)	(0.353)	(0.0520)	(0.0511)	(0.599)	(0.573)	(0.225)	(0.202)
lnGDP	-6.353	-6.549	53.815	60.706	-290.964*	-217.063	-47.685	-48.347
	(74.85)	(74.97)	(73.49)	(72.62)	(174.4)	(162.4)	(33.00)	(33.09)
gas price no taxes	240.449***	240.23***						
	(81.66)	(81.56)						
gas tax	-3.763	-3.752						
	(25.22)	(25.21)						
gas price			-16.821	-17.865				
			(19.50)	(19.66)				
oil per capita	7211.473	7235.793	-7884.717	-9278.706				
	(10175.8)	(10120.7)	(17329.9)	(17305.9)				
gas per capita			-1413.198	-1245.722				
			(2200.3)	(2181.3)				
coal per capita			-10.85*	-12.21*				
			(6.385)	(6.566)				
electric price			0.099	0.107				
			(0.121)	(0.122)				
grow_elec					-5.258	-3.77	-0.872	-0.802
					(5.688)	(5.039)	(0.920)	(0.929)
% hydro					0.629	1.627	1.043	0.98
					(2.065)	(1.788)	(0.718)	(0.737)
% nuclear					-4.819*	-4.592*	-0.499	-0.421
					(2.493)	(2.454)	(0.678)	(0.681)
FIT wind							9.827	12.061
							(25.86)	(26.07)
FIT pv					18.874	39.523		
					(60.28)	(53.99)		
REC levels					15.092**	13.524**	2.133	2.102
					(7.234)	(6.352)	(1.376)	(1.352)

Table 7 – Share of articles receiving NPL citations

	Biofuels		*Solar*		*Wind*	
	N	% with NPL citation	N	% with NPL citation	N	% with NPL citation
USA						
1991-1995	168.8	4.1%	618.0	4.2%	56.3	8.9%
1996-2000	318.1	7.8%	669.3	6.7%	96.3	4.2%
2001-2005	408.9	4.9%	981.4	4.4%	163.4	6.1%
2006-2011	3532.9	0.7%	4055.2	0.6%	600.6	0.3%
Total	4428.6	1.7%	6323.9	2.2%	916.5	2.3%
foreign						
1991-1995	630.2	0.6%	2054.0	2.8%	264.7	3.0%
1996-2000	959.9	1.2%	3669.7	4.2%	406.8	3.2%
2001-2005	1783.1	1.1%	6172.6	2.3%	758.7	2.2%
2006-2011	12847.1	0.2%	23055.8	0.2%	3268.4	0.2%
Total	16220.4	0.4%	34952.1	1.2%	4698.5	0.9%

Note: Table uses weighted shares, weighting each article by the share of U.S. and foreign authors

Table 8 – Time to first NPL citation hazard regression

	full sample			*articles from 2000 or later*		
	Biofuels	Solar	Wind	Biofuels	Solar	Wind
citation lag	0.706***	0.333***	0.531***	1.359***	0.691***	0.973**
	(0.142)	(0.0500)	(0.157)	(0.326)	(0.109)	(0.480)
(citation lag)^2	-0.0249***	-0.0167***	-0.0191***	-0.0682***	-0.0412***	-0.0562
	(0.00754)	(0.00331)	(0.00731)	(0.0231)	(0.00866)	(0.0377)
Multiple country dummy	-0.0301	-0.0294	-1.234**	-0.0478	-0.0821	-1.775*
	(0.282)	(0.123)	(0.626)	(0.322)	(0.151)	(1.058)
Constant	-8.987***	-7.500***	-7.353***	-11.22***	-8.621***	-8.283***
	(0.948)	(0.597)	(1.096)	(1.324)	(0.659)	(1.655)
N	55790	164957	21895	35568	95725	13670
AIC	1407.3	4306.9	831.2	881.6	2444.9	487.3
BIC	3827.1	6930.5	2709.8	2170.4	3780.1	1525.5
log likelihood	-432.6	-1891.5	-180.6	-288.8	-1081.5	-105.6

* p<0.10, ** p<0.05, *** p<0.01
Standard errors in parentheses. All standard errors clustered by article.
All regressions include publication year x country fixed effects, with publications from 2001 in the US as the base category.

Table 9 -- Tests for diminishing returns to government R&D

	Biofuels	Energy Efficiency	Solar Energy	Wind Energy	Biofuels	Energy Efficiency	Solar Energy	Wind Energy
Cumulative effects:								
R&D	0.168	0.093	-1.386*	0.975	1.39***	0.21***	1.694	0.518*
	(0.840)	(0.155)	(0.799)	(0.623)	(0.439)	(0.0614)	(1.066)	(0.275)
R&Dsq	0.0189***	0.0003	0.0113***	-0.0062				
	(0.00670)	(0.000237)	(0.00314)	(0.00843)				
R&Dx100%					3.461***	0.177	-0.358	1.715***
					(1.189)	(0.400)	(2.725)	(0.437)
N	300	280	280	280	300	280	280	280
AIC	2263.2	2171.6	2542.5	1730.7	2300.5	2101.9	2583.7	1717.7
BIC	2655.8	2636.9	2891.5	2108.7	2693.1	2567.2	2932.6	2095.7
F RD 1st stage	162.3	237.4	68.98	114.9	40.91	33.11	4.411	45.95
F RDsq 1st stage	3043.6	1876.7	452.6	777.4				
F RDx100% 1st stage					42.3	22.55	56.3	30.25
Hansen J p-value	0.143	0.0135	0.0632	0.263	0.254	0.232	0.257	0.491
Endog. test p-value	0.689	0.527	0.477	0.845	0.917	0.0557	0.415	0.262

Robust standard errors with correction for autocorrelation in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table 10 –Does the quality of publications change?

Biofuels	(1) Biofuels	(2) Energy Efficiency	(3) Solar	(4) Wind	(5) Biofuels	(6) Energy Efficiency	(7) Solar	(8) Wind
ln(publications, cited year)	-0.170**	-0.186*	0.0321	-0.167*	-0.0351	-0.0451	-0.220	-0.263*
	(0.0537)	(0.0894)	(0.0774)	(0.0784)	(0.0803)	(0.121)	(0.153)	(0.112)
ln(publications, cited year)^2					-0.0136*	-0.0312	0.0300*	0.0221
					(0.00667)	(0.0174)	(0.0132)	(0.0177)
multicountry	0.228***	0.220***	0.167***	0.238***	0.227***	0.218***	0.168***	0.238***
	(0.0347)	(0.0595)	(0.0337)	(0.0645)	(0.0347)	(0.0596)	(0.0335)	(0.0644)
cited year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	13347	10163	27697	4232	13347	10163	27697	4232
log likelihood	-32424.9	-22919.8	-74139.3	-8534.9	-32421.9	-22917.0	-74128.9	-8534.2

Dependent variable is total citations received by each article after publication. For articles with authors from multiple countries, each country/article observation is weighted by the share that country in the total number of authors.

Standard error, clustered by article, in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Appendix A: Web of Science Keyword Searches

Our publication data were provided as a custom database from Thomson Reuters, created using keyword searches in Web of Science developed by the researchers in consultation with experts at Thomson Reuters. As noted in the text, we devised searches that would be narrow, so as to avoid irrelevant articles, as such articles would respond differently to alternative energy R&D trends and thus bias our results downward. Thus, for each category, we devised keyword searches that are refined by focusing on specific subject categories in the Web of Science database

Biofuels

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

Energy Efficiency

Note: The energy efficiency search strategy uses the three separate searches below, and excludes publications included in other energy categories,

#1	TS = ("waste heat" NEAR/3 (recover* OR convert* OR utilize* OR use)) OR TS = ("district heat*" OR "district cool*") OR TS = (("LED" OR "light emitting diode") NEAR/1 (lighting OR lightbulb* OR "light bulb*" OR lamp* OR “solid state light*” OR “solid state lamp*”)) OR TS = (("CFL" OR "compact fluorescent") NEAR/1 (lighting OR lightbulb* OR "light bulb*" OR lamp*)) OR TS = "solid state light*"
#2	(TS = ((electric OR hybrid OR “electric drive” OR “hybrid drive”) NEAR/0 (vehicle* OR automobile* OR "auto" OR "autos" OR"car" OR "cars")) NOT TS = (“fuel cell*” OR “fuel-cell*” OR “Hybrid Car-Parrinello”)) Refined by: [excluding] Web of Science Categories=(ASTRONOMY ASTROPHYSICS OR DERMATOLOGY OR INFECTIOUS DISEASES OR LITERATURE BRITISH ISLES OR MEDICINE GENERAL INTERNAL OR PHARMACOLOGY PHARMACY OR HISTORY PHILOSOPHY OF SCIENCE OR SPECTROSCOPY OR VETERINARY SCIENCES OR CLINICAL NEUROLOGY) Timespan=1991-2010. Databases=SCI-EXPANDED. Lemmatization=Off

#3	(TS=((energy OR fuel OR gas* OR electric* OR petrol*) NEAR/1 (consum* OR use OR using OR usage OR burn*) NEAR/1 (reduc* OR less OR lower OR decreas*)) OR TS=((energy OR fuel OR gas* OR petrol) NEAR/1 (efficien* OR economy OR mileage OR productivity) NEAR/1 (improv* OR increas* OR better OR greater OR more)) OR TS=((energy OR fuel OR gas* OR electric* OR petrol*) NEAR/1 (saving* OR save OR saves OR saved))) NOT TS = ("fuel cell*" OR “fuel-cell*”) Refined by: [excluding] Web of Science Categories=(ENVIRONMENTAL STUDIES OR FOOD SCIENCE TECHNOLOGY OR ZOOLOGY OR NUTRITION DIETETICS OR PHYSIOLOGY OR BIOCHEMISTRY MOLECULAR BIOLOGY OR BIOLOGY OR METEOROLOGY ATMOSPHERIC SCIENCES OR CARDIAC CARDIOVASCULAR SYSTEMS OR ASTRONOMY ASTROPHYSICS OR MATHEMATICS APPLIED OR ROBOTICS) AND Web of Science Categories=(ENERGY FUELS OR ENGINEERING CHEMICAL OR ENGINEERING ELECTRICAL ELECTRONIC OR ENGINEERING MECHANICAL OR THERMODYNAMICS OR ENVIRONMENTAL SCIENCES OR CONSTRUCTION BUILDING TECHNOLOGY OR MATERIALS SCIENCE MULTIDISCIPLINARY OR ENGINEERING ENVIRONMENTAL OR METALLURGY METALLURGICAL ENGINEERING OR TELECOMMUNICATIONS OR COMPUTER SCIENCE INFORMATION SYSTEMS OR COMPUTER SCIENCE HARDWARE ARCHITECTURE OR COMPUTER SCIENCE THEORY METHODS OR TRANSPORTATION SCIENCE TECHNOLOGY OR ENGINEERING MULTIDISCIPLINARY OR MATERIALS SCIENCE PAPER WOOD OR AGRICULTURAL ENGINEERING OR AUTOMATION CONTROL SYSTEMS OR MATERIALS SCIENCE CERAMICS OR COMPUTER SCIENCE SOFTWARE ENGINEERING OR MINING MINERAL PROCESSING OR ENGINEERING INDUSTRIAL OR COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE OR AGRONOMY OR MATERIALS SCIENCE COMPOSITES OR MATERIALS SCIENCE TEXTILES OR OPERATIONS RESEARCH MANAGEMENT SCIENCE OR ENGINEERING MARINE OR ENGINEERING OCEAN OR URBAN STUDIES) Databases=SCI-EXPANDED Timespan=1991-2010 Lemmatization=Off

Solar Energy

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

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

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

Wind Energy

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

Appendix B: Results without European Union fixed effects

To account for unreported energy R&D support from the European Union to its member countries, the results in the text include separate year fixed effects for EU countries. Table B1 of this appendix compares the results from the first-differenced instrumental variables model with and without the EU- specific fixed effects for the R&D coefficients, and Table B2 compares the results for the controls.

Dechezlepretre and Popp (2015) report that the EU's share of renewable energy R&D funding is small, so that unaccounted for EU spending is unlikely to be a major factor in scientific productivity. For example, EU RD&D investments dedicated to the six technologies covered by the EU Strategic Energy Technology Plan (wind, solar (photovoltaics and concentrated solar power), electricity grids, bioenergy, carbon capture and storage, fuel cells and hydrogen and nuclear fission) accounted for just 11 percent of public R&D spending on these technologies in 2010. Tables B1 and B2 confirm that the effect of any omitted EU-wide R&D is minimal, as the results are virtually the same when excluding the EU-specific year effects. The cumulative effects of energy R&D range from 6 to 15 percent higher without the EU-specific year trends, which is consistent with the expected bias, assuming that EU funding follows similar trends as country-level funding. F-tests for the joint significance of all EU-specific year fixed effects fail to reject the null hypothesis for all technologies except wind. Nonetheless, to avoid any bias from unreported EU-wide energy R&D spending, the results in the main text include EU-specific year fixed effects.

Table B1 – Effect of EU-specific year effects on first-differenced IV panel regression results: Government R&D

	Biofuels		Energy Efficiency		Solar Energy		Wind Energy	
	EU x year	no EU	EU x year	no EU	EU x year	no EU	EU x year	no EU
RD	0.152***	0.159***	0.0197***	0.0197***	-0.0267	0.00768	0.0983**	0.0956**
	(0.0153)	(0.0178)	(0.00649)	(0.00688)	(0.276)	(0.313)	(0.0440)	(0.0469)
RD(t-1)	0.376***	0.361***	0.0460***	0.0442***	0.601***	0.609***	0.197***	0.196***
	(0.0259)	(0.0296)	(0.00675)	(0.00764)	(0.167)	(0.216)	(0.0520)	(0.0571)
RD(t-2)	0.429***	0.397***	0.0722***	0.0732***	0.483**	0.566*	0.335***	0.338***
	(0.0196)	(0.0229)	(0.00886)	(0.00965)	(0.246)	(0.301)	(0.0547)	(0.0635)
RD(t-3)	0.00719	0.0726	-0.0128	-0.0153	0.154	0.141	0.0313	0.0412
	(0.0719)	(0.0789)	(0.0296)	(0.0323)	(0.144)	(0.172)	(0.0580)	(0.0613)
RD(t-4)	0.176**	0.294**	0.0141	0.0105	0.447**	0.390*	0.0263	0.0222
	(0.0891)	(0.121)	(0.0253)	(0.0273)	(0.191)	(0.220)	(0.0585)	(0.0662)
RD(t-5)	-0.101	-0.0270	-0.0303	-0.0132	0.0350	0.136	0.0130	0.00634
	(0.109)	(0.118)	(0.0255)	(0.0284)	(0.243)	(0.303)	(0.0532)	(0.0594)
RD(t-6)	0.0940	0.139	0.0167	0.0149	0.385**	0.374*	0.0618	0.0942
	(0.104)	(0.127)	(0.0176)	(0.0182)	(0.194)	(0.219)	(0.0519)	(0.0596)
RD(t-7)	0.0585	0.0587	0.000165	0.0136			0.121***	0.145***
	(0.0965)	(0.121)	(0.0229)	(0.0246)			(0.0451)	(0.0510)
RD(t-8)	0.0884	0.00444	-0.0329	-0.0355				
	(0.118)	(0.147)	(0.0255)	(0.0280)				
RD(t-9)	0.558***	0.559***	0.0381	0.0426				
	(0.160)	(0.209)	(0.0268)	(0.0318)				
RD(t-10)	0.318**	0.365*	0.0598***	0.0646***				
	(0.152)	(0.214)	(0.0211)	(0.0192)				
Cumulative effects:								
R&D	2.155***	2.383***	0.191***	0.219***	2.078***	2.223***	0.883***	0.939***
	(0.348)	(0.444)	(0.0520)	(0.0535)	(0.599)	(0.727)	(0.225)	(0.246)
N	300	300	280	280	280	280	280	280
AIC	2351.3	2367.0	2110.7	2099.1	2676.4	2682.1	1766.5	1760.1
EU x year p-value	0.284		0.458		0.155		0.0220	
F RD 1st stage	32.51	29.82	18.01	20.31	11.40	13.44	38.49	32.07
Hansen J p-value	0.293	0.411	0.183	0.197	0.491	0.632	0.502	0.540
Endog. test p-value	0.979	0.737	0.240	0.267	0.0459	0.124	0.258	0.192

Standard errors in parentheses. All models use robust standard errors with correction for autocorrelation.
*: significant at 10% level. **: significant at 5% level. ***: Significant at 1% level.

Table B2 – Effect of EU-specific year effects on first-differenced IV panel regression results: Controls

	Biofuels		Energy Efficiency		Solar Energy		Wind Energy	
	IV	exog	IV	exog	IV	exog	IV	exog
Cumulative effects:								
R&D	2.155***	2.383***	0.191***	0.219***	2.078***	2.223***	0.883***	0.939***
	(0.348)	(0.444)	(0.0520)	(0.0535)	(0.599)	(0.727)	(0.225)	(0.246)
lnGDP	-6.353	-41.66	53.815	64.658	-290.964*	-231.318	-47.685	-40.879
	(74.85)	(87.24)	(73.49)	(71.41)	(174.4)	(175.7)	(33.00)	(34.85)
gas price no taxes	240.449***	293.545***						
	(81.66)	(94.19)						
gas tax	-3.763	-8.82						
	(25.22)	(28.14)						
gas price			-16.821	-10.405				
			(19.50)	(21.33)				
oil per capita	7211.473	5833.289	-7884.717	9947.448				
	(10175.8)	(8473.9)	(17329.9)	(13063.0)				
gas per capita			-1413.198	-1353.485				
			(2200.3)	(2104.4)				
coal per capita			-10.85*	-7.408				
			(6.385)	(7.198)				
electric price			0.099	0.108				
			(0.121)	(0.108)				
grow_elec					-5.258	2.215	-0.872	-0.033
					(5.688)	(4.959)	(0.920)	(0.888)
% hydro					0.629	1.102	1.043	1.236
					(2.065)	(2.335)	(0.718)	(0.794)
% nuclear					-4.819*	-1.843	-0.499	-0.163
					(2.493)	(2.551)	(0.678)	(0.677)
FIT wind							9.827	16.495
							(25.86)	(27.57)
FIT pv					18.874	-70.981		
					(60.28)	(66.22)		
REC levels					15.092**	7.073	2.133	1.753
					(7.234)	(7.325)	(1.376)	(1.440)

Standard errors in parentheses. All models use robust standard errors with correction for autocorrelation.
*: significant at 10% level. **: significant at 5% level. ***: Significant at 1% level.

Appendix C: Determining the Optimal Lag Length

As noted in the text, we employ several strategies to identify the appropriate lag length. First, as is common in the literature, the primary criterion is finding the minimum AIC statistic across a range of models (e.g. Crespi and Geuna, 2008).[19] One complication is that our model includes not only lagged values of the R&D variables, but also of several control variables. These controls are often individually insignificant. Thus, we also run models including only a single year of the control and policy variables, to see if this changes the recommended number of lags. We initially examine models including up to ten lags of R&D. Adding additional lags is problematic when the model includes renewable energy policy variables as only two countries in our sample adopted renewable energy certificates before 2003, so that estimates of the lagged value of the REC variable beyond eight years are unreliable. However, adding up to 11 years of lags is possible for biofuels and energy efficiency.[20]

In addition, even collinearity among the R&D variables themselves may cause the AIC statistic to favor smaller lags. Thus, we also run both the full and single control models using a polynomial distributed lag model (PDL). The PDL models provide similar results to the main first differenced models, but by requiring fewer parameters to estimate multiple lags, they offer the potential for lower standard errors on the R&D coefficients. Table C1 presents the AIC statistics.

In some cases, these strategies recommend different lag lengths. Thus, we also consider the cumulative long-run effect of R&D when evaluating the lag length. These long-run effects,

[19] Alternatively, we also calculated the BIC statistic for each model, which includes a greater penalty for including irrelevant variables. As a result, collinearity among the lagged R&D values often leads the BIC to recommend fewer lags than the AIC. However, in the case of collinear lagged R&D values, we can still estimate long-run effects that are jointly significant, even when individual year coefficients are estimated imprecisely. Moreover, leaving out relevant lags would lead to omitted variable bias. Thus, I focus on the findings of the AIC statistic in the discussion that follows.

[20] The first US states to adopt REC limits do so in 1998, and Italy adopts an REC limit in 2002. Most states first limits appear in the data in 2003.

illustrated in Figure 4, tend to be similar across the various estimation techniques, so we focus on the results using first differenced instrumental variables. When the AIC statistic conflicts across models, we use the pattern of cumulative effects as an additional guide. In particular, we consider the lag length at which the cumulative effect of government energy R&D spending levels out, as this is evidence that all relevant lags have been included.

Turning first to biofuels, we see relatively consistent recommendations across models. Using the AIC criterion we find an optimal lag length of ten years using the full model and eleven years using a single year of controls. The PDL results suggest an optimal lag length of nine years in the full model and ten when using a single year of controls. As the AIC is lowest for the full FD model, which has an optimal lag of ten years, I use a lag length of **ten** years for analyzing the biofuels results.

The results for wind demonstrate the potential advantages of the PDL model and for looking at the results of models with a single year of controls. Using the standard first differenced model, the AIC suggests an optimal lag length of just two years. Using just a single year of controls, the AIC suggests using seven years of lagged data. In contrast, the PDL model suggests ten year lags in the full model and four lags when using a single year of controls. This difference, however is not caused by adding lags of energy R&D, but rather lags of the tradable permit policy variable, which becomes large and significant beginning with the ninth lagged variable. However, beyond nine years of lags, this coefficient is identified off of just one country, as only the United States had renewable credit trading before 2002. Thus, we turn to the cumulative effect of energy R&D to help sort through the options. In all four models (FD, FD single control, PDL, PDL single control), the cumulative effect grows significantly between years 6 and 7 before leveling off, suggesting a lag of seven years (as suggested by the FD model with a single year of controls) is appropriate. Indeed, the eighth and ninth lags of energy R&D are insignificant and close to zero. Thus, I use a lag length of **seven** years for analyzing the wind energy results.

For solar energy, shorter lags appear to suffice, although the choice using the AIC statistic varies across models. In the FD model, the AIC statistic is minimized with just four lags in the full model and with nine lags when using just a single year of controls. In contrast, the AIC suggests ten years of lags in the PDL model with full controls, and eight when using just a single year of controls.

Thus, again we turn to the estimated cumulative effects for more information. Here, we see a large increase in the cumulative effect between years three and four, and again between years five and six. While the individual coefficient for year *t*-6 is only significant at the 10 percent level, both the large size of the coefficient, as well as the continued growth in the cumulative effect through year six, suggests that **six** years should be included in the solar energy models. There is little value to adding additional lags of energy R&D in the PDL models. Indeed, results in year 10 seem to be largely driven by changes in the REC variable, as the sign of the effect of R&D becomes negative in the model with full controls. In contrast, the cumulative effect of energy R&D remains level beyond year 6 in the model using a single year of control variables.

The appropriate lag length is more difficult to identify in the case of energy efficiency, as the effect of energy R&D itself is smaller. In the FD model, the AIC statistic suggest a lag of two years using the full model but ten years when using a single year of controls. In the PDL model, the AIC is minimized using a lag of five years using a full set of controls, and four lagged years with a single year of controls. Given these divergent results, we turn to the cumulative effect of energy R&D. The cumulative effect levels off after year six before rising again beginning in year

9. Among models with 9 to 11 lags, the AIC is minimized in the main FD model when using 10 lags. Thus, I use **ten** years of lagged R&D when evaluating energy efficiency.

Table C1 – AIC statistic for alternative number of years of lagged R&D

Biofuels

	FDIV		PDL IV	
	all controls	single year	all controls[c]	single year[b]
AIC lag 1:	2811.9	2806.3		
AIC lag 2:	2406.5	2400.5		
AIC lag 3:	2408	2402.4		
AIC lag 4:	2401.2	2401.9	2401.2	2435.8
AIC lag 5:	2397.2	2403.8	2425.6	2434.8
AIC lag 6:	2401.6	2399	2425.6	2428.4
AIC lag 7:	2394.1	2400.7	2421.9	2428.5
AIC lag 8:	2378.3	2400.2	2413.7	2427.5
AIC lag 9:	2356	2372.6	**2388.8**	2417.5
AIC lag 10:	**2351.3**	2356.9	2389.8	**2397.2**
AIC lag 11:	2351.5	**2352.3**	2408.2	2420.1

Energy Efficiency

	FDIV		PDL IV	
	all controls	single year	all controls[a]	single year[c]
AIC lag 1:	2183.4	2173		
AIC lag 2:	**2060.2**	2045.4		
AIC lag 3:	2071.4	2047.3		
AIC lag 4:	2079.9	2047.8	2076.7	**2047.8**
AIC lag 5:	2083.9	2044.7	**2071.9**	2054.0
AIC lag 6:	2091.5	2046.8	2086.6	2054.4
AIC lag 7:	2098	2049	2096.9	2062.5
AIC lag 8:	2106.5	2049.2	2093.6	2069.7
AIC lag 9:	2114.7	2048.4	2114.4	2067.9
AIC lag 10:	2110.7	**2038.6**	2130.9	2062.4
AIC lag 11:	2114.6	2039.2	2133.7	2063.4

Solar Energy

	FDIV		PDL IV	
	all controls	single year	all controls[b]	single year[a]
AIC lag 1:	2742.8	2732.5		
AIC lag 2:	2704.1	2688.2		
AIC lag 3:	2714.2	2689.9		
AIC lag 4:	**2672.9**	2641.2	2652.6	2632.9
AIC lag 5:	2674.6	2643.2	2657.3	2636.6
AIC lag 6:	2676.4	2640.4	2654.2	2625.5
AIC lag 7:	2676.8	2641	2644.7	2624.6
AIC lag 8:	2682.4	2640.3	2638.9	**2623.7**
AIC lag 9:	2689.1	**2636.9**	2631.8	2626.0
AIC lag 10:	2760.2	2639.4	**2623.8**	2626.0

Wind Energy

	FDIV		PDL IV	
	all controls	single year	all controls[b]	single year[c]
AIC lag 1:	1839.5	1844.9		
AIC lag 2:	**1738.7**	1742.5		
AIC lag 3:	1748.1	1744		
AIC lag 4:	1750.8	1745	1764.3	**1745.0**
AIC lag 5:	1756.5	1746.9	1771.8	1758.3
AIC lag 6:	1763.3	1747.8	1773.2	1765.3
AIC lag 7:	1766.5	**1737.3**	1770.6	1760.5
AIC lag 8:	1777.1	1739.3	1788.2	1767.7
AIC lag 9:	1747.7	1740.2	1758.6	1768.7
AIC lag 10:	1746.3	1737.7	**1741.3**	1768.7

a: 2nd degree polynomial; b: 3rd degree polynomial; c: 4th degree polynomial

Appendix D: Polynomial Distributed Lag Results

The following tables present results using the polynomial distributed lag (PDL) model. To accommodate instrumental variables for current energy R&D, I used the following steps:

1. Estimate a first-stage model for R&D using the same number of lagged controls as in the final PDL regression to be estimated.
2. Obtain predicted energy R&D from this regression.
3. Use the predicted current R&D, along with actual values of lagged R&D, to construct first differenced data.
4. Use first differenced data to construct the polynomials for estimation. Polynomials from degree 2-4 were used, as were lags from 4-10 years.

As noted in the text, there are few differences between the PDL results and those simply including all lags in the model. Moreover, there is little gain in efficiency from using the PDL model, but it does hele to identify the appropriate lag length for each technology. Thus, except for the discussion of lag length, I focus on the unrestricted model in the discussion in the text. For reference, complete PDL results are presented here. Table D1 shows the results for government R&D for both individual years and the cumulative effect, and Table D2 shows the cumulative effects of the various controls included in each model.

The main differences between the unrestricted results and the PDL results are as follows. First, for biofuels, the FD results suggest a cyclical pattern, with strong effects also found in years two and four in the first differenced model before picking up again in year *t*-9. In contrast, the PDL model suggests a more gradual effect of energy R&D, with the effect initially peaking in year *t*-2, and gradually fading until recurring in years nine and ten. Similarly, for energy efficiency, the largest single year impact does not occur until year *t*-10 in the PDL mode, compared to year *t*-2 in the FD model. However, the long-run cumulative effects are similar in both models.

Table D1 – Polynomial distributed lag regression results: Government R&D

	Biofuels		Energy Efficiency		Solar Energy		Wind Energy	
	IV	exog	IV	exog	IV	exog	IV	exog
RD	0.1435***	0.1449***	0.0152**	0.0100	0.2678**	0.3909***	0.1085*	0.1276***
	(0.0123)	(0.0143)	(0.0067)	(0.0073)	(0.1228)	(0.0852)	(0.0568)	(0.0323)
RD(t-1)	0.3121***	0.3127***	0.0440***	0.0410***	0.3872***	0.4460***	0.2507***	0.2647***
	(0.0278)	(0.0283)	(0.0045)	(0.0053)	(0.1041)	(0.1013)	(0.0451)	(0.0411)
RD(t-2)	0.3822***	0.3826***	0.0523***	0.0506***	0.4532***	0.4692***	0.2658***	0.2764***
	(0.0181)	(0.0181)	(0.0085)	(0.0084)	(0.1431)	(0.1405)	(0.0514)	(0.0551)
RD(t-3)	0.3768***	0.3774***	0.0457***	0.0447***	0.4655***	0.4604***	0.2021***	0.2104***
	(0.0366)	(0.0365)	(0.0104)	(0.0100)	(0.1620)	(0.1549)	(0.0555)	(0.0604)
RD(t-4)	0.3210***	0.3219***	0.0298***	0.0291***	0.4243***	0.4198***	0.1078*	0.1147*
	(0.0614)	(0.0615)	(0.0107)	(0.0104)	(0.1518)	(0.1461)	(0.0552)	(0.0594)
RD(t-5)	0.2415***	0.2426***	0.0102	0.0095	0.3295**	0.3472**	0.0312	0.0370
	(0.0729)	(0.0736)	(0.0106)	(0.0104)	(0.1355)	(0.1419)	(0.0497)	(0.0523)
RD(t-6)	0.1669**	0.1681**	-0.0074	-0.0083	0.1811	0.2428	0.0204	0.0252
	(0.0720)	(0.0737)	(0.0104)	(0.0103)	(0.1797)	(0.1966)	(0.0397)	(0.0397)
RD(t-7)	0.1274*	0.1287*	-0.0175*	-0.0183*			0.1237**	0.1271**
	(0.0717)	(0.0744)	(0.0102)	(0.0101)			(0.0517)	(0.0497)
RD(t-8)	0.1553*	0.1564*	-0.0145	-0.0149				
	(0.0855)	(0.0884)	(0.0103)	(0.0102)				
RD(t-9)	0.2843**	0.2854**	0.0073	0.0078				
	(0.1175)	(0.1195)	(0.0135)	(0.0134)				
RD(t-10)	0.5502***	0.5515***	0.0535**	0.0556**				
	(0.1971)	(0.1978)	(0.0231)	(0.0230)				
R&D	3.0612***	3.0723***	0.2186***	0.2067***	2.5085***	2.7764***	1.1101***	1.1830***
	(0.4413)	(0.4611)	(0.0627)	(0.0633)	(0.6621)	(0.7313)	(0.2765)	(0.2792)
N	300	300	280	280	280	280	280	280
AIC	2389.8	2389.8	2085.2	2084.1	2641.1	2637.1	1770.6	1770.0
BIC	2630.5	2630.5	2332.3	2331.3	2862.8	2858.8	2017.8	2017.1

Standard errors in parentheses. All models use robust standard errors with correction for autocorrelation.
*: significant at 10% level. **: significant at 5% level. ***: Significant at 1% level.

Table D2 – Polynomial distributed lag regression results: Controls

	Biofuels		Energy Efficiency		Solar Energy		Wind Energy	
	IV	exog	IV	exog	IV	exog	IV	exog
Cumulative effects:								
R&D	3.0612***	3.0723***	0.2186***	0.2067***	2.5085***	2.7764***	1.1101***	1.1830***
	(0.4413)	(0.4611)	(0.0627)	(0.0633)	(0.6621)	(0.7313)	(0.2765)	(0.2792)
lnGDP	-2.1057	-2.8775	53.2447	51.4496	-1.9e+02	-1.7e+02	-41.2015	-41.6738
	(82.2929)	(81.7642)	(69.5854)	(69.2378)	(169.4939)	(168.1542)	(33.6773)	(33.8759)
gas price no taxes	221.8810***	221.2094***						
	(84.4800)	(83.8467)						
gas tax	-3.1941	-3.1393						
	(27.2391)	(27.1973)						
gas price			-5.2159	-6.8539				
			(21.2337)	(21.2315)				
oil per capita	7.1e+03	7.1e+03	-1.2e+04	-1.2e+04				
	(1.1e+04)	(1.1e+04)	(8.5e+03)	(8.4e+03)				
gas per capita			-1.4e+03	-1.5e+03				
			(2.2e+03)	(2.2e+03)				
coal per capita			-11.4431**	-12.0952**				
			(5.2265)	(5.2584)				
electric price			0.0753	0.0774				
			(0.1234)	(0.1238)				
grow_elec					-3.2878	-2.7047	-0.7717	-0.7167
					(5.1424)	(1.0026)	(1.0169)	(1.0026)
% hydro					-0.0743	0.0766	0.4473	0.4144
					(1.6134)	(1.5613)	(0.8803)	(0.8923)
% nuclear					-3.6867	-3.6852	-0.0177	0.0323
					(2.5755)	(2.5097)	(0.7553)	(0.7685)
FIT wind							25.0600	25.8369
							(26.4234)	(26.5270)
FIT pv					55.4779	55.7739		
					(60.0910)	(59.4518)		
REC levels					14.4206**	13.5407*	2.6813*	2.6451*
					(7.2394)	(6.9293)	(1.5560)	(1.5268)

Standard errors in parentheses. All models use robust standard errors with correction for autocorrelation.
*: significant at 10% level. **: significant at 5% level. ***: Significant at 1% level.

Appendix E: Detailed methodology for NPL citation calculations

The increased probability of an NPL citation resulting from an additional one million dollars of new energy R&D funding depends on both the number of articles induced by this R&D each year and on the probability of an article from any given year being cited in the future. Thus, Figures 6 and 7 use the results of both our R&D estimation and the NPL citation regression:

(1) $Q_{i,t} = \sum_{s=0}^{T} \beta_{t-s} R_{i,t-s} + \sum_{s=0}^{T} \boldsymbol{\gamma}_{\mathbf{t-s}} \mathbf{POLICY}_{i,t-s} + \sum_{s=0}^{T} \boldsymbol{\delta}_{\mathbf{t-s}} \mathbf{X}_{i,t-s} + \alpha_i + \eta_t + \epsilon_{i,t}$

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

Let $\boldsymbol{\beta} = [\beta_t \; \beta_{t-1} \; \cdots \; \beta_{t-T}]$ represent a row vector of coefficients on contemporary and lagged R&D from equation (1). Using the results of (2), the annual probability that a publication from year *t* receives a citation in year *t* + s can be written as:

$$w_s = \exp\{\alpha_0 + \alpha_1 s + \alpha_2 s^2\}$$

Next, define a matrix $\mathbf{W_{Annual}}$ representing the annual probability that an article published in year *t*, represented by the rows of the matrix, will be cited in year *t* + *s*, represented by the columns of the matrix:

$$\mathbf{W_{Annual}} = \begin{bmatrix} w_0 & w_1 & w_2 & \cdots & w_M \\ & w_0 & w_1 & & w_{M-1} \\ & & w_0 & & \\ & & & \ddots & \\ & & & & w_0 \end{bmatrix}$$

Similarly, the cumulative probability of a publication in year *t* receiving a citation by year *t* + *s* is given by the following matrix:

$$\mathbf{W_{Cumulative}} = \begin{bmatrix} w_0 & w_0 + w_1 & w_2 + w_1 + w_0 & \cdots & w_M + w_{M-1} + \ldots + w_0 \\ & w_0 & w_1 + w_0 & & w_{M-1} + w_{M-2} + \ldots + w_0 \\ & & w_0 & & \\ & & & \ddots & \\ & & & & w_0 \end{bmatrix}$$

Using these matrices, the product $\boldsymbol{\beta}\mathbf{W}_{\mathbf{Annual}}$ yields a row matrix with the increase in the annual probability of an NPL citation each year after an additional one million dollars of energy R&D, and the product $\boldsymbol{\beta}\mathbf{W}_{\mathbf{Cumulative}}$ yields a row matrix with the increase in the cumulative probability resulting from an additional one million dollars energy R&D.