NBER WORKING PAPER SERIES

GOVERNMENT FUNDING AND THE DIRECTION OF ACADEMIC ENERGY RESEARCH

David Popp
Myriam Gregoire-Zawilski
Lizhen Liang
Daniel Acuna

Working Paper 34856
http://www.nber.org/papers/w34856

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2026

Financial support for this research was provided by the Alfred P. Sloan Foundation, grant G-2020-12618. We also acknowledge a scientific research grant from Dimensions that provided access to data used in this paper. We thank Sarah Armitage, Hugo Jales, and participants in the NBER Science of Science Funding conference for helpful comments. Ziqiao Chen, Jingni Zhang, and Hanlin Zhang provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Government Funding and the Direction of Academic Energy Research
David Popp, Myriam Gregoire-Zawilski, Lizhen Liang, and Daniel Acuna
NBER Working Paper No. 34856
February 2026
JEL No. O38, Q48, Q55

ABSTRACT

Does government funding influence the choice of research topics? Novel grant-making modalities such as the Advanced Research Projects Agency-Energy (ARPA-E) program aim to encourage scientists to take on difficult-to-solve, wicked societal problems such as clean energy. Yet little causal evidence exists linking funding and research direction, with most existing studies focusing on health sciences. We provide new evidence on the effect of funding on clean energy research, addressing two questions: (1) Do scientists change the focus of their research in response to targeted government funding opportunities? (2) If so, what types of calls for funding best attract new researchers? Using data on grants from the Department of Energy and National Science Foundation, we combine text and regression analyses to compare the publication trajectories of funded scientists to a set of matched controls. After funding, the research of funded scientists becomes more similar to the grant topic than that of the matched controls. The effect is largest for ARPA-E, which explicitly aims to attract new scientists to clean energy research, suggesting that agency efforts to attract new researchers to a topic area can succeed. General calls for funding such as offered by traditional NSF directorates generate less movement.

David Popp
Syracuse University
The Maxwell School
Department of Public Administration
and NBER
dcpopp@syr.edu

Myriam Gregoire-Zawilski
University of Ottawa
Graduate School of Public and
International Affairs,
and Institute of the Environment
mgregoir2@uottawa.ca

Lizhen Liang
Syracuse University
lliang06@syr.edu

Daniel Acuna
University of Colorado Boulder
Daniel.Acuna@colorado.edu

Science funding agencies face increasing pressure to demonstrate the relevance and effectiveness of their investments. Since the early 2000s, funders have experimented with novel grant-making modalities for encouraging scientists to take on difficult-to-solve, wicked societal problems. Funders have supported high-risk projects in emerging areas such as food, climate and health sciences. In the energy space, governments have increasingly delegated responsibility for setting targeted research priorities aligned with sustainability goals to energy agencies (Mandt et al., 2020; Meckling et al., 2022). In the United States, the Department of Energy began piloting this approach in 2009 with the Advanced Research Projects Agency-Energy (ARPA-E) program, modeled on the Department of Defense's pioneering DARPA program. From its inception, the goals of ARPA-E explicitly listed supporting high-risk/high-reward research to transform the energy landscape, including through creating new communities of researchers "not traditionally... involved in the topic area" (National Academies of Sciences, Engineering, and Medicine, 2017, p. 112). Yet we lack causal evidence on whether mission-oriented, high-risk funding programs induce lasting changes in scientists' research directions rather than simply deepening existing activity.

To be truly impactful, government investments intending to address wicked problems should not simply substitute for other sources of funding available, but convene scientists around society's important emerging problems that are underfunded. Yet most academic literature evaluating government research only asks whether funding increases research output (Jacob & Lefgren, 2011; Rosenbloom et al., 2015). While studies focusing on narrow fields of research typically find a positive relationship between funding and the quantity of research output (Blume-Kohout, 2012; Popp, 2016), existing literature says little about the mechanisms by which research increases. Does it come from moving existing researchers into the field? Does it encourage junior researchers to enter the field? Moreover, are there differences across types of funding calls? Are mission-oriented programs such as ARPA-E more successful bringing new researchers into the field? Notably few papers in the scientific policy literature address the direction of research, and fewer still attempt to make causal links between funding and research direction. Rare exceptions focus on health sciences (Azoulay et al., 2011; Myers, 2020).

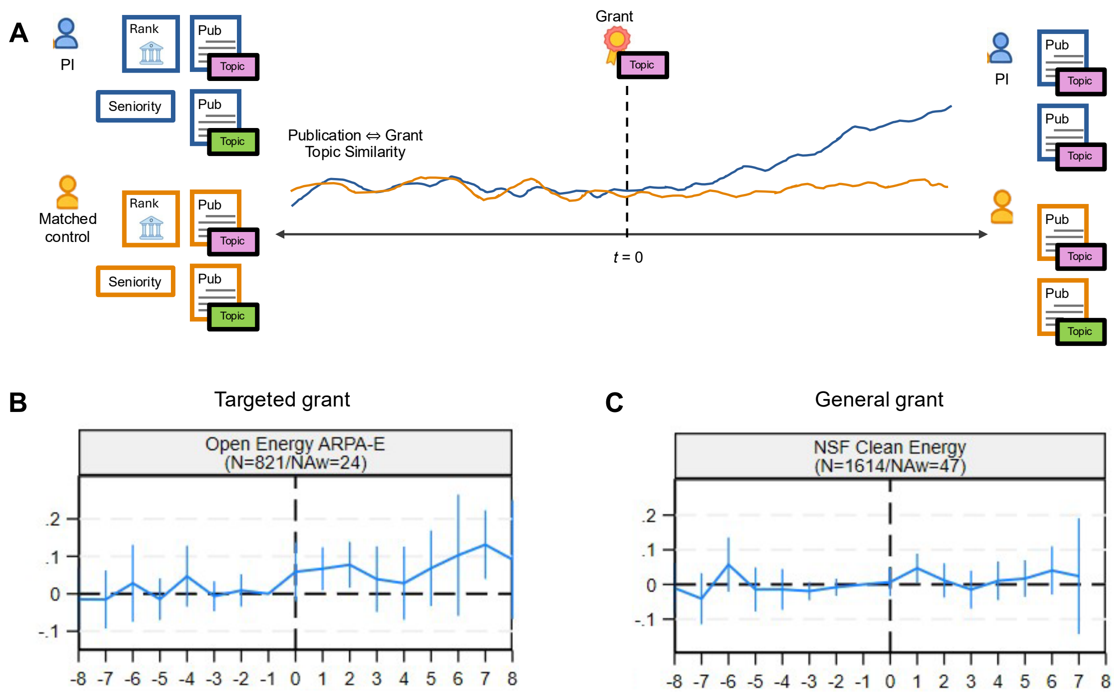


Figure 1 | Using a matched sample to quantify effect of funding on research topics. **a.** For each PI we identify a matched control with similar characteristics (e.g. career length, institution quality) in the three years prior to funding. Prior to funding, both the PI and matched control publish papers on similar topics (purple and green above). In the illustrated example, after receiving funding related to the purple topic, the PI publishes more papers on that topic. Their work becomes more similar to the funded topic than their matched control. **b.** Event study estimating differences in similarity to the grant of PI publications relative to controls. Funding from targeted programs such as ARPA-E moves researchers closer to their grant topic. **c.** Funding from general open-ended programs has no effect on research topics.

Here, we examine research funding from the US Department of Energy (DOE) and National Science Foundation (NSF). While our primary focus is on clean energy research funding, to assess the effectiveness of different funding models we include both mission-driven programs such as ARPA-E and related programs featuring open-ended calls for proposals, such as DOE's Office of Science and select NSF programs. Open-ended programs allow for a broad range of proposed ideas, typically bounded by a specific discipline or broad research area, potentially lowering the incentive to switch research topics. We combine data on grant abstracts from these programs with publication data of funded researchers and a matched set of controls. Using topic modeling, we compare the set of topics in each scientist's research portfolio before and after funding to assess whether funding moves a research portfolio closer to the funded topic. Both the

PIs and their matched controls have similar characteristics and work on similar research prior to funding. (Figure 1a). While all funding has at least some effect on topic choice, consistent with its mission, the largest changes in research direction come from ARPA-E funding (Figure 1b & c). Using patent citations to measure impact, we show that ARPA-E funded research by scientists moving into the energy space is highly cited.

As the ARPA model has since been extended to both other fields (e.g. ARPA-H for health) and other countries (the United Kingdom's Advanced Research Innovation Agency), our findings have broad applicability across a range of scientific fields. Moreover, our research demonstrates how rich textual data accessible through bibliometric sources can identify matched controls. These methods could be applied to other areas of science policy, such as using patent records to study the effectiveness of policies designed to support small businesses, making our methodology broadly applicable in science policy analysis.

I. Literature Review

Our research contributes to two strands of literature: (1) energy R&D policy research, and (2) research on the direction of scientific activity. Regarding energy R&D policy, few papers provide causal evidence on the impact of publicly funded clean energy R&D, such as research conducted in government labs or funded by the government through grants and subsidies. Most studies simply include aggregated public R&D expenditures as one of several variables in more general studies of the drivers of clean energy patenting (e.g., Gugler et al., 2024; Johnstone et al., 2010; Verdolini & Galeotti, 2011). Focusing directly on public energy research, Popp (2016) links data on scientific publications to public energy R&D funding, finding that $1 million 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. Goldstein and Narayanamurti (2018) and a 2017 National Academies of Science report (National Academies of Sciences, Engineering, and Medicine, 2017) provide early assessments of ARPA-E, but their authors acknowledge that their analyses are only preliminary and not causal in nature. With the initial ARPA-E awards now over 15 years old, we provide a more thorough causal evaluation of this and other programs' long-term effects on the scientific trajectories of funded researchers.

Most academic literature evaluating government research asks whether funding increases research output, but does not study the mechanisms by which research increases. Studies asking whether funding increases the level of research activity find mixed results (e.g., Jacob & Lefgren, 2011; Rosenbloom et al., 2015), with those focusing on narrow fields of research more likely to find a positive relationship between funding and research outputs (e.g., Blume-Kohout, 2012 on NIH; Popp, 2016 on energy R&D). But this is not the appropriate question to ask if government funding aims to redirect research. Even if researchers are able to substitute funding from multiple sources to maintain a given level of research activity, government funding may affect the topics they chose to pursue.

A large literature addresses why scientists choose the topics they do (Stephan, 2010), including scientific priority (e.g., Merton, 1957), reputation, and financial rewards (e.g., Stephan, 1996). While large shifts in research topics over the course of one's career are rare (Jia et al., 2017), sudden shocks can compel researchers to shift their focus onto new topics. Using the COVID-19 pandemic as a natural experiment, Hill et al. (2025) find that researchers who made the largest pivots in response to these pressing problems published in lower-impact journals. The high cost of pivoting indicates that governments can play a role to support and de-risk the pursuit of new research directions aligned with the societal priorities of funding agencies.

Fewer papers investigate the effects of government policies on the direction of research directly. Azoulay et al. (2011) document how the nature of different research grants affects the riskiness of topics chosen, but do not consider how funding affects the choice of research field itself. Two recent papers address topic choice more directly. Myers (2020) compares applications to open-ended and targeted NIH competitions. Researchers are indifferent between a one standard deviation redirection to less similar science and about $1 million additional grant funds. Mancuso and Broström (2026) focus on mission-oriented grants, comparing the similarity between scientist's publications and funding calls from the Swedish Foundation for Strategic Research. They construct three control groups, comparing (1) applicants to non-applicants, (2) non-winning applicants to non-applicants, and (3) winning applicants to non-winning applicants. While they use matching to construct these control groups, they do not use text analysis for matching. Instead, they use journal subject categories to identify researchers working in similar fields. Both successful and non-successful applicants move their research closer to the funding announcement than non-applicants, suggesting funding calls themselves influence research direction.

Our research is unique in that it uses topic modeling to analyze text and create a matched sample of funded and non-funded researchers. Topic modeling has been shown to be suitable for producing transparent and interpretable textual features for science of science (Kozlowski et al., 2019). Previous work has emphasized the advantage of topic models in providing structured, human-readable themes compared to recent natural language processing techniques such as deep learning (Chai & Li, 2019; Zhao et al., 2021). While some studies use machine learning to devise novel measures of research outputs, such as the aforementioned Mancuso and Broström (2026) paper or Furman and Teodoridis' (2020) research showing that lowering of cost of motion sensing research encouraged new entrants into the field, none use machine learning to construct a matched sample. Instead, studies using matching to identify a control group working in similar fields rely on journal names or keywords (e.g., Mancuso & Broström, 2026; Shin et al., 2022). Modern topic modeling methods allow for a more detailed classification of research themes, capturing deeper semantic structures beyond simple keyword matching (e.g., Rijcken et al., 2022). This approach improves the interpretability of matched samples by revealing underlying research trajectories rather than relying only on surface-level metadata. Our proposed methods tap into rich textual data accessible through bibliometric sources and state-of-the-art machine learning techniques to identify matches that have similar research topics and trajectories as the funded researchers in our sample. These methods could be applied to other areas of science policy, such as using patent records to study the effectiveness of policies designed to support small businesses. Given the growing importance of interpretability in text-based research, integrating topic modeling with deep learning could provide an optimal balance between performance and transparency, making our methodology broadly applicable in science policy analysis.

II. Research Setting

Our analysis compares programs relevant for clean energy at the U.S. Department of Energy and the National Science Foundation. We include *general* programs that solicit research on a range of topics, *energy-specific* programs soliciting research on clean energy topics defined by researchers, and *targeted* energy programs where researchers respond to specific funding requirements (Table 1).

General	DOE, Office of Science Basic Energy Sciences *NSF General:* • NSF, Chemical Catalysis (except NSF Clean Energy) • NSF, Electronic & Photonic Materials (except NSF Clean Energy)
Open Energy-specific	*ARPA-E Open Energy-Specific* • DOE ARPA-E Open • DOE ARPA-E IDEAS *NSF Open Energy-Specific* • NSF Clean Energy Technology (awards made by NSF Chemical Catalysis or Electronic& Photonic Materials) • NSF Energy and Environment • NSF Energy for Sustainability • NSF Sustainable Energy Pathways
Targeted Energy-specific	DOE ARPA-E Focus DOE EERE: Bioenergy Technologies DOE EERE: Solar Energy Technologies DOE EERE: Wind Energy Technologies DOE EERE: Water Power Technologies

Table 1 | Categorization of research programs

We consider three programs at DOE spanning this range of funding options. ARPA-E began in 2009. ARPA-E aims to identify and fund high-risk, potentially high return research in advanced alternative energy technologies (National Academies of Sciences, Engineering, and Medicine, 2017). It solicits research proposals through both Focused Funding Opportunity Announcements (FOA) targeting specific technology areas and tri-annual Open FOAs to ensure they do not miss potential innovative opportunities in other areas. Because ARPA-E was designed to bridge the gap between basic and applied research at DOE (Goldstein & Narayanamurti, 2018), we include additional DOE programs on either side of this spectrum. The DOE's Office of Science Basic Energy Sciences program supports basic research using open solicitations. DOE's Office of Energy Efficiency and Renewable Energy (EERE) supports applied clean energy research. Like ARPA-E, EERE uses targeted FOAs to solicit research proposals that specify detailed technological improvements to be achieved. Its goal is to address gaps in technology development where the private sector will not invest (Office of Energy Efficiency and Renewable Energy, n.d.).

NSF standard grants include multiple directorates overseeing programs supporting research in a broad range of science and engineering fields through open calls for proposals. We study two programs commonly used as funding sources by researchers also funded by DOE: the Chemical Catalysis and Electronic & Photonic Materials Programs. Both programs attract researchers working on energy, but also fund related research on other topics. This provides an opportunity to see if broad programs can attract scientists interested in switching topics. Furthermore, in select years, NSF made additional funds available for clean energy across multiple directorates. Proposals were to be submitted to the appropriate NSF disciplinary area. We separately identify such awards made by these two directorates.

While most NSF programs use open call for proposals, a few programs specifically solicit research pertaining to clean energy. These NSF programs solicit proposals for a range of clean energy technologies. As the audience of potential grantees at NSF is likely larger than DOE, understanding the impact of such programs at NSF is important, as they might bring in researchers that are new to the broad field of energy research. We include three such programs (Table 1). Energy and the Environment awards are through the Industry–University Cooperative Research Centers (IUCRC) program. The Energy for Sustainability program supported fundamental engineering research to enable sustainable production of electricity and fuels. Finally, unlike the aforementioned energy programs in NSF, which are administered in a single directorate, Sustainable Energy Pathways involved five different NSF directorates and required interdisciplinary teams of researchers. Appendix A describes these programs in greater detail.

III. Data and Sample

We collect data on 2,598 grants awarded by DOE or NSF between 2009 and 2016, which we combine with publication data for each award's PI from the Dimensions database. Dimensions contains data on 146 million scientific publications, integrated from repositories like Crossref and arXiv, and from 130 academic publishers. Among academic databases, Dimensions high levels of reliability and completeness (Delgado-Quirós & Ortega, 2024). We use Dimensions to collect information on the publication portfolios of the funded researchers and their potential matches. Separately, we collect data on (1) scientists funded by the programs in our study (Table 1) and (2) the abstracts of the funded projects.

Data on Department of Energy grants were collected using various strategies, depending on the availability of information for each program. A list of all Principal Investigators funded by ARPA-E, their grant numbers and the related funding opportunity announcement identifier were obtained from the Department of Energy through a Freedom of Information Request. We also collected grant information on four sub-EERE programs on solar energy (SETO), wind energy (WETO), waterpower (WPTO) and bioenergy (BETO) from the USAspending.gov website. However, this source only makes grant numbers and recipient institutions available. We completed manual searches to further identify principal investigators (PIs) funded by EERE from various technical reports published on the OSTI.gov and on the DOE's website. We then cross-referenced this information using data about funding acknowledgments contained in these authors' publications. Information on awards made by the Office of Basic Energy Sciences program was retrieved from the Office of Science Portfolio Analysis and Management Systems (PAMS). The PAMS data contains the abstract of many of the funded projects. We manually searched for missing abstracts from this and other programs, on the DOE website and in reports of project summaries. For the National Science Foundation awards, we downloaded data from the Federal Reporter database, filtering for our sampled programs. This database contains detailed information on the award number, PI, recipient institution, and grant abstract.

Having identified which grants were funded through the different funding opportunity announcements published by the different programs, we linked information on the PIs to their respective grants across the different programs. In total, we begin with 2,697 grants with start dates between the years 2009-2016. To track the portfolios of these PIs over time, we manually match the names of PIs funded by these awards to researcher names in Dimensions to identify unique researcher identification numbers in Dimensions for each. We dropped awards for which a unique identifier could not be determined (e.g. because of records split across different institutions, or two researchers with similar names and research portfolios), leaving us with 2,598 potential awards with 1,965 unique PIs for matching.

IV. Research Topics & Similarity

To measure changes in research direction in response to funding, we first train a topic model to represent the range of topics covered in a publication. We then calculate the similarity between each researcher's publications in year *t*, and the funded research topic, represented by the grant abstract, by computing the cosine distance between their respective representations. Our training data includes both publications of our PIs and a random sample of potentially matched researchers (see section V).

There are multiple ways to compute similarities. We use the text of the grant's abstract and the abstract of each scientist's publications as the primary data to estimate similarities. We use Natural Language Processing (NLP) techniques to find *semantic* similarity between sets of documents based on their raw text. We avoid direct bag-of-words comparison because raw term-frequency or TF-IDF vectors can exceed hundreds of thousands of dimensions, and distance metrics lose power in such sparse spaces (Aggarwal et al., 2001). Early studies addressed this problem by representing each document with Latent Semantic Analysis (Deerwester et al., 1990). We go further and represent each document with Latent Dirichlet Allocation topic proportions (Blei et al., 2003). This topic-based representation delivers faster and more reliable similarity estimates than raw TF-IDF or bag-of-words baselines (Achakulvisut et al., 2016; Manning et al., 2008). Additionally, topic-based representations are more transparent and can be directly validated compared to more recent deep neural embeddings such as BERT (Devlin et al., 2019). Finally, topic models run orders of magnitude faster than deep learning models and handle full-length documents naturally. In sum, our approach balances speed, accuracy, and transparency for estimating similarities.

To validate the topic model and illustrate which research areas PIs come from, Figures 2 and 3 show the top topics by program, ranked by average appearance across papers funded by each program. Appendix Table B1 shows the ten most probable terms in these topics. Topics with an applied energy or environmental focus are highlighted in green. Validating our topic model, blue bars show that the top topics across papers funded by each program relate to that program's objectives (e.g. photovoltaics, thin films, and power systems in EERE, catalysts and materials in NSF Other General).

Differences in the topic distributions of funded papers and papers published in the three years before funding illustrate movement across different research topics. Green bars show the distribution of paper topics for PIs whose work is most similar to their grant, while red bars show these topics for PIs least similar to their grant.

We see smaller differences in the pre-and post-funding distribution of topics for general programs than in those targeting energy. We highlight ARPA-E as an example. The top three topics in published papers are energy applications (energy systems, photovoltaics, fuel processing, or energy storage). Except for energy storage in Open ARPA-E, these topics are shared by PIs whose prior record is similar to their grant (green bars). The top topics of dissimilar PIs include scientific methods relevant for clean energy, such as thermal material properties (relevant for solar) and genomics (relevant for biofuels). Relative to funded papers, however, the papers of these PIs are less likely to include topics directly related to energy (except for batteries and energy storage). For these PIs, ARPA-E funding supports researchers with the necessary scientific skills but with broader interests besides energy. In contrast, the main topics in NSF Clean Energy are not energy-specific, but rather pertain to the goals of the NSF programs included in our sample (catalysts and materials). There is less difference between the topics of the PIs previous research and the funded work, suggesting NSF Clean Energy grants were not attracting new scientists to energy research.

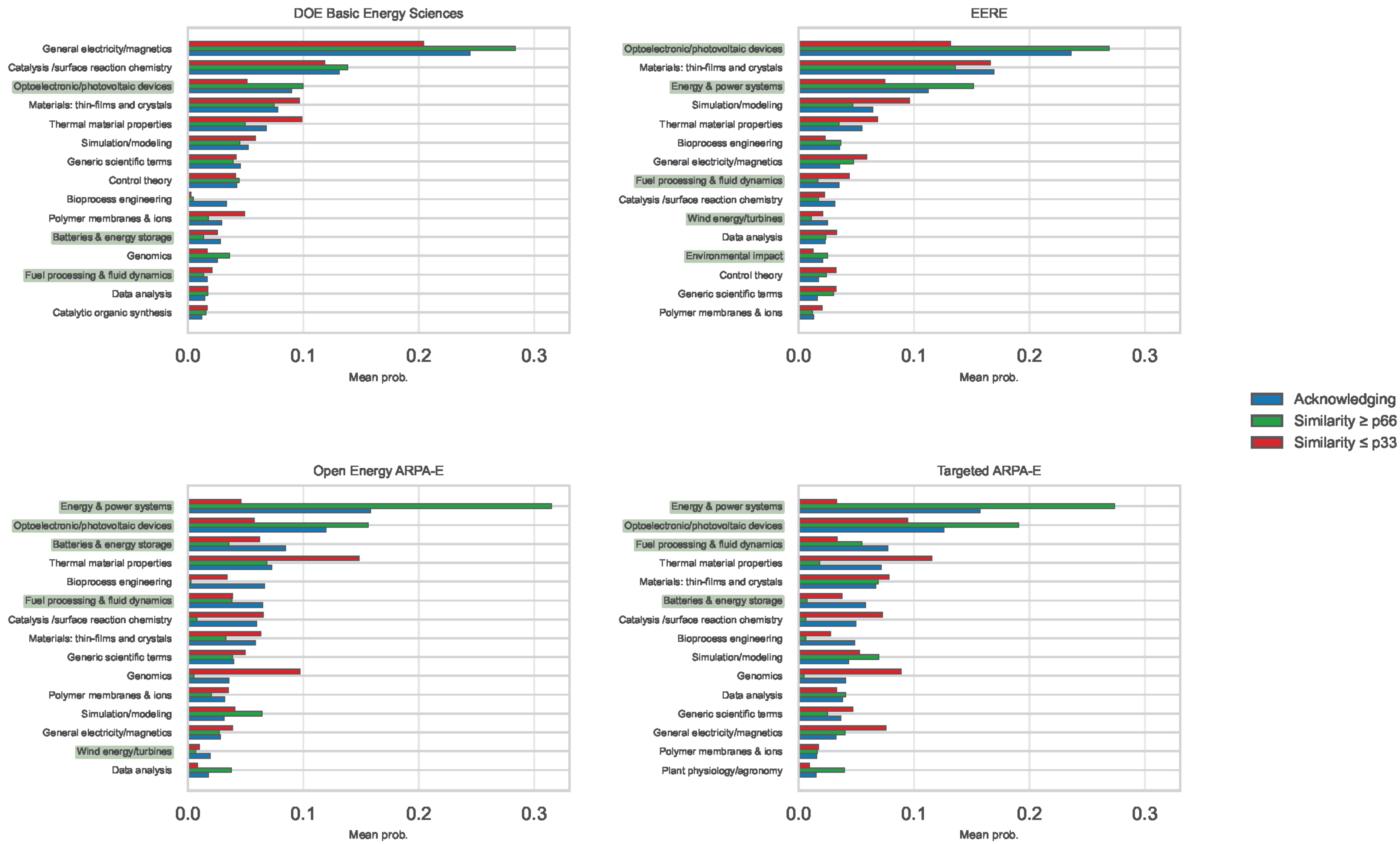


Figure 2 | Top topics by Program, Department of Energy. Figure shows the top topics by their average appearance across papers funded by each program (blue bars). Red and green bars show the distribution of paper topics for the same PIs in all papers published in the three years prior to funding. Green bars represent PIs whose prior work is most similar to the grant, while red bars represent PIs whose prior work is least similar. Differences in the topic distributions of funded papers and papers published in the three years before funding illustrate movement across different research topics. Topic names highlighted in green relate to energy or environmental applications.

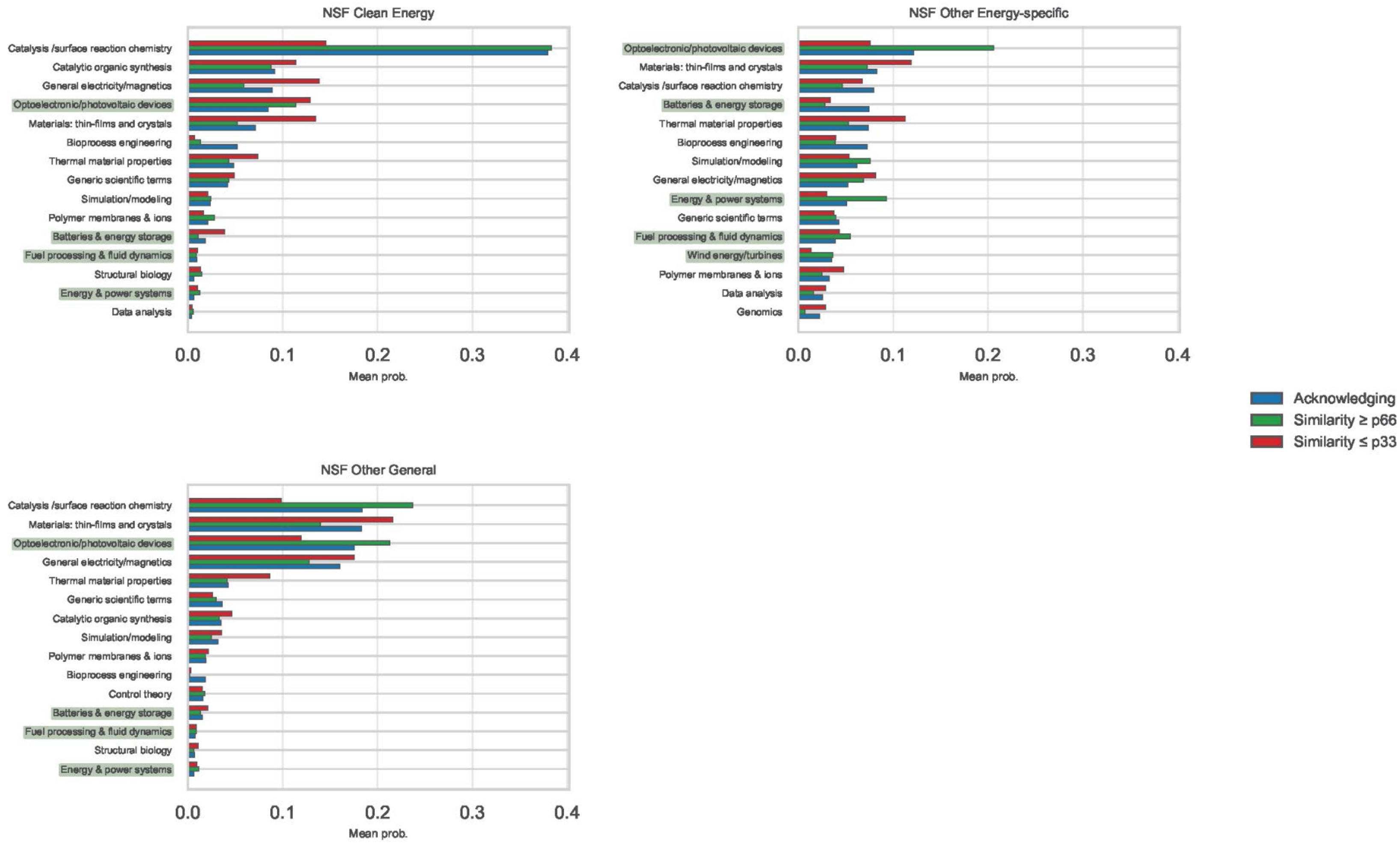


Figure 3 | Top topics by Program, National Science Foundation. Figure shows the top topics by their average appearance across papers funded by each program (blue bars). Red and green bars show the distribution of paper topics for the same PIs in all papers published in the three years prior to funding. Green bars represent PIs whose prior work is most similar to the grant, while red bars represent PIs whose prior work is least similar. Differences in the topic distributions of funded papers and papers published in the three years before funding illustrate movement across different research topics. Topic names highlighted in green relate to energy or environmental applications. NSF Other General includes all awards made by NSF general programs except those under the NSF Clean Energy initiative.

As additional validation, in Appendix B we perform a qualitative analysis of how grants cluster in topic space using a principal component analysis (PCA) (Appendix Figure B1). For both agencies, the PCA reveals meaningful clusters. Within NSF, the two main programs cluster as they should: Catalysis/surface reaction chemistry is the highest loading topic for the Chemical Catalysis program, and Optoelectronic/photovoltaic devices the highest loading topic for the Electronic and Photonic Materials program (X-axis). Along the Y-axis, awards pertaining to energy and power systems predominately go to the NSF Open Energy programs. Similarly, within the Department of Energy, we see clusters dividing basic and applied energy. Energy and power systems are the highest loading topic for ARPA-E and EERE. Both programs focus on energy applications. In contrast, for DOE's Basic Energy Sciences, which funds basic sciences related to energy, general electricity/magnetics is prominent. Along the y-axis, EERE solar awards cluster in the lower right (Optoelectronic/photovoltaic devices), while bioenergy is in the upper right (Catalysis/surface reaction chemistry).

A. Similarity to Research Grants

To measure changes in research direction, we calculate the similarity between each researcher's publications in year *t*, and the funded research topic, represented by the grant abstract, by computing the cosine distance between their respective representations. Measuring similarities using NLP involves two steps: (1) transforming the sets of documents into vector representations, and (2) using an appropriate distance function for estimating their similarities. Ideally, the vector representation should aim to capture the semantics of the documents using a low dimensional space. Some representations such as vectors holding the counts of unigrams or *n*-grams are very high dimensional, causing several difficulties (Aggarwal et al., 2001). For this reason, dimensionality reduction techniques are performed first (Manning et al., 2008), as described in Appendix B. As for the choice of the distance function, the cosine similarity is preferred because it allows comparison of documents of different lengths. Using dimensionality reduction and cosine similarity, documents can be compared effectively. Appendix B discusses the calculation of cosine similarity in more detail.

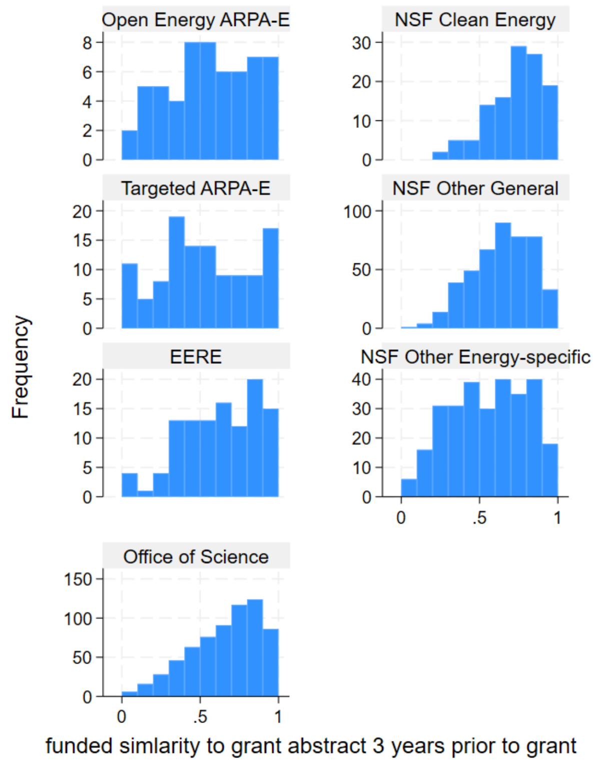


Figure 4 | Distribution of similarity to funded research. Figure shows the distribution of the similarity between a PI's publications in the three years prior to the start of their grant and the grant abstract. Similarity ranges from 0 (completely different) to 1 (exactly the same). While most distributions are skewed towards 1, both ARPA-E programs have less skewed distributions.

Figure 4 shows the distribution of the similarity of a PI's research to their grant abstract in the three years prior to funding. In general, the distribution is skewed, with most PIs doing research fairly similar to the work that gets funded. However, ARPA-E brings in PIs with greater dispersion of similarity, including several whose work is very dissimilar to the grant abstract (e.g. on the left of each distribution). EERE and NSF's other energy-specific programs are also less skewed than the general programs, but they do not bring in as many distant researchers as ARPA-E.

V. Creating a Matched Sample

While these figures reveal differences between ARPA-E and the other programs in the topical focus of the PIs' prior research, attributing changes post-funding to the program design poses challenges. PIs more inclined to make large pivots may differ from other researchers. For example, ARPA-E PIs tend to have more experience and publications than other PIs in our sample (SI Figure G3). To establish causal links between funding and research direction, we need to know whether the funded scientists pursued research directions considerably different from what they would have done in the absence of the grant. This requires a control group of scientists not funded by these programs to account for other external factors affecting the choices of both funded and non-funded scientists. The ideal counterfactual are scientists who were not funded by these research programs, but who worked on similar research topics as our PIs, making them just as likely to have pursued the same research directions at the time of the grant. We identify such a control group of authors who are similar to funded scientists in research interest, career stage, and other features *up to the time in which our treated scientists start receiving funding*. We then compare the research profiles of our funded scientists and matched controls after the start of the funded scientist's grant (Figure 5).

Our matching algorithm chooses the smallest distance between the PI and their proposed match for the variables described below. These distances consider both the similarity between the publication portfolios of the PI and proposed match, as well as the distance between each researcher's portfolio and the funded research project (measured using the funded grant abstract). Including similarity of both the PI and their proposed match to the funded project is important, as researchers work on a range of topics and we want to ensure similarity on research topics targeted by the grant. Consider a scientist who, pre-treatment, equally divides her time between two topics, A and B. The scientist then receives a grant pertaining to topic B. Researchers dividing their time equally between topics A and C or between topics B and C would have portfolios of similar distance to the PI. But since the first never worked on the grant topic, they would be less likely to receive funding on a grant related to topic B than the PI. By including similarity to the funded research, both the PI and matched control can be considered equally plausible candidates to do the work in the grant. Thus, we calculate two different types of similarity:

- $grantdist_{i,t}$ is the cosine similarity between each researcher's publications in year t, and the funded research topic, which is represented by the grant abstract. This variable is the dependent variable in our analysis, and values in the years prior to funding are used in the matching process, as described below.

- *pairdist*$_3$ is the cosine similarity between the PI and each potential match's publications in the 3 years prior to funding. This distance is also used for matching.

Each similarity measure varies from zero (the two sets are completely different) to one (the two sets are exactly the same).

Because PIs may be moving towards the grant topic even before funding, we identify potential matches using the weighted average of annual distances between the funded research and prior publications in the 3 years leading to the grant. This ensures that we select matches whose publications portfolios were moving toward the topics of the funded research at a similar pace as the PIs, while also accounting for heterogeneity in publishing rates across these years. This measure, *WAdist*$_3$. identifies sets of PIs and controls whose research is similarly distant to the funded topic in years prior to the grant. *WAdist*$_3$ is defined as the square root of the sum of the squared distance between the funded research topic and the publications of a PI and the match:

$$WAdist_3 = \sqrt{\sum_{s=1}^{3} w_{t-s}\left(grantdist_{PI,t-s} - grantdist_{M,t-s}\right)^2}$$

where

$$w_{t-s} = \frac{N_{PI,t-s} + N_{M,t-s}}{\sum_{s=1}^{3}(N_{PI,t-s} + N_{M,t-s})}$$

is the weight placed on each year, $grantdist_{i,t-s}$ is the distance between each researcher's publications in year *t-s*, and the funded research topic, and $N_{i,t-s}$ is the number of publications for each researcher in year *t-s*. *WAdist*$_3$ therefore captures how similarly distant the PIs and their potential matches are from the topics of the funded research.

A. Identifying the Control Set

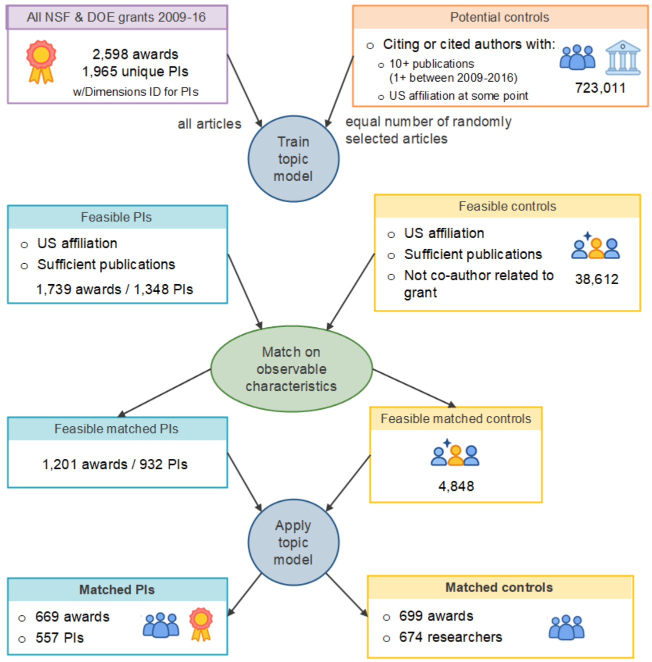


Figure 5 | Creating a Matched Sample.

We follow three steps to efficiently identify potential control scientists. First, we obtain publication data on a set of *potential controls* who either cite or are cited by our funded researchers. To ensure that potential researchers represented realistic grant competitors, we imposed eligibility criteria focusing on researchers who authored at least 10 research publications, have a U.S. affiliation at some point in their career, and had at least one publication between 2009 and 2016, ensuring active researchers during the funding period. As discussed in more detail in Appendix B, we train the topic model using the full set of articles published by our funded researchers and an equal number of randomly selected articles from the set of potential controls. Second, we further narrow this set of potential controls to *feasible controls* by identifying controls with similar observable characteristics. These include research productivity, measured by both number of publications (quantity) and number of citations (quality), number of co-authors, a measure of the breadth of topics in prior publications, years of experience (calculated based on the date of the author's first publication in Dimensions), and the quality of an author's research institution. The resources institutions provide affect research productivity. While we don't want to compare only researchers from the same university, we can select as controls researchers working in similar environments. Third, we use topic modeling to identify *matched controls*, obtaining at most a single match for each funded researcher whose research is both similar overall to the PI (*pairdist*$_3$) and is a similar distance from the funded research (*WAdist*$_3$). We briefly describe this process below. Figure 5 summarizes the matching process and shows the number of awards and researchers kept after each step. A fuller description of the matching process is provided in Appendix C.

To generate our final set of PIs and controls, we drop both PIs and controls with insufficient publications. This includes having no publications in any given year during a window 3 years prior to and 5 years after the start of the grant. We also require both PIs and potential controls to have a US academic affiliation during this period. As we do not know if the grants have co-PIs or other recipients, we exclude potential matches who are co-authors on any papers acknowledging the grant. If no papers acknowledge the grant, we exclude co-authors of the PI on any paper published within 5 years of the start of the grant. In the analysis comparing research to the grant abstract, we drop awards with abstracts of 50 words or less. This gives us a set of *feasible PIs and controls* (Figure 5) We then apply the following filters to both PIs and their potential matches to create a set of *feasible matched PIs and controls*:

- We require that both the PI and match must be at a US academic institution of similar quality, where quality is measured using the Carnegie classification for higher education. If they have multiple affiliations in their publication record, a US academic institution must be listed over 50% of the time.
- Defining experience as the number of years since a researcher's first publication, the experience of the PI and potential match at the time of the grant must be +/- 50% of each other.
- To ensure similar breadth of research portfolios, the entropy of the PI and potential match in the three years prior to the grant must be +/- 25%. Entropy captures the breadth of topics on which a scientist publishes, with higher values indicating more breadth. Among two

researchers with comparable similarity measures to a given paper topic, a researcher with additional experience in other areas may be able to combine the knowledge from each field in novel ways. The calculation of entropy is described in Appendix B.

Finally, using the topic model results, we combine $pairdist_3$ and $WAdist_3$, choosing for each PI the potential match that produces the lowest (1-$pairdist_3$) + $WAdist_3$. This ensures we select matches that are closest to the PIs in both the topics of their research portfolios and their closeness to the funded research. To ensure that the match is a quality match, we only keep potential matches where $(1\text{-}pairdist_3)+WAdist_3 \leq 0.15$ and $WAdist_3 \leq 0.10$. These thresholds were chosen through manual inspection of a set of PIs and matched controls to determine levels of similarity at which matched researchers no longer serve as reasonable matches for the PI. Appendix D provides examples of the closest and furthest matches.

Our main results are based on a single matching rule, but to assess sensitivity to our choices we present results using other matching rules in the appendix. Each alternate matching rule differs based on how similar the *level* and *quality* of each researcher's publication record must be to be considered a feasible match. Our main matching rule restricts both the number of publications and citations in the 3 years prior to the grant to be within +/- 50%. In choosing between matching rules, we trade off generating closer matches on publication quantity and quality against reductions in sample size. In the appendix, we present alternate results with a larger sample by not restricting on citation counts, and on a smaller sample by limiting publication counts to be within +/- 25%

A. Evaluation of Matching

Prior to funding, both the PIs and matched controls have similar research profiles, becoming more similar to the grant topic even before funding is granted (Figure 6). Plotting by calendar year (Appendix Figure D1), the research of both PIs and controls becomes more similar to the grant projects until 2013. Similar trends have been documented in other studies of clean energy innovations, including in patents (Popp et al., 2022, 2024), and venture capital (van den Heuvel & Popp, 2023). Clean energy innovation experienced a boom period in the early 2000s, peaking between 2010 and 2012. While innovation on some clean energy topics, such as electric vehicles, has increased again, for technologies such as solar PV and wind, most innovation metrics have yet to return to the peak of the early 2010s boom. Trends for PIs and their matched controls diverge after funding. Matching overall trends in clean energy research, the research topics of our controls begin moving away from the topics of the grants in our sample, with the largest decline in open-energy-specific and targeted programs (Appendix Figure D1). In contrast, while PIs do not continue to move closer to their grant topic, funding maintains their attention on related topics longer.

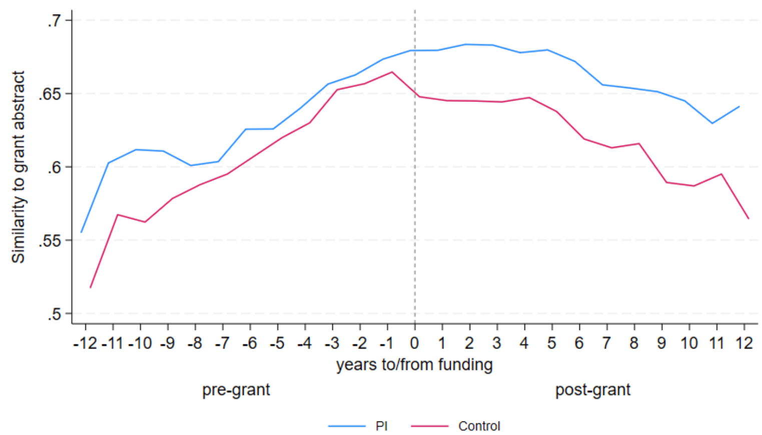


Figure 6 | Average Similarity to Grant Abstract, years to/from funding. The research topics of both PIs (blue) and their matched controls (red) become more similar to the PI's grant abstract in the years prior to funding. Their paths diverge after funding.

Illustrating the ability of our matching procedure to identify a viable set of controls, Appendix Figure D2 shows almost no relationship between similarity between PIs and the full set of feasible controls. Appendix Table D1 compares the means of key variables for the PIs and controls in our sample. Appendix Figure D3 shows the full distribution of each variable. To check for balance in the matched sample, we present the standardized difference of means (SMD) and variance ratio (VR).[1] For an appropriately balanced sample, Rubin suggests SMD < 0.25 and VR between 0.5 and 2 (Rubin, 2001). Our main matching rule keeps 699 awards. The research of both PIs and controls is similarly distanced from the grant abstract (0.683 vs 0.675). Each have between 8 and 9 publications per year. While PIs publish slightly

[1] Here $SMD = \frac{\bar{x}_{PI}-\bar{X}_{Control}}{\sqrt{(s_{PI}^2-s_{Control}^2)/2}}$ and $VR = {s_{PI}^2}/{s_{Control}^2}$, where $\bar{X}$ represents the sample mean and s^2 the sample variance.

more, are cited a bit more frequently, and have slightly more co-authors, all are within the acceptable limits of an SMD less than 0.25 and VR less than 2. While limiting to publication counts within +/- 25% narrows the gap on publication counts, it does so at a cost of dropping an additional 167 awards. Similarly, not considering citations at all increases the sample size to 934 awards, but results in a VR greater than 2 for both the number of citations and co-authors (Appendix Tables D2 & D3). To compare the effect of matching on the sample across programs, Appendix Figure G3 shows boxplots of PI characteristics for both matched PIs and PIs in the full sample for each funding program.

VI. Model Specification

We now turn to estimating the causal effect of research funding on changes in research direction. Figure 6 shows small differences in the level of similarity to the grant abstract in the years directly prior to funding. To remove any remaining differences between the PIs and controls, we use both event study and difference-in-differences regressions. Individual fixed effects control for any remaining time-invariant differences between PIs and their matched controls. This normalization allows us to compare PIs and controls with the small pre-treatment differences removed. The event study allows us to verify that PIs and their matched controls were on similar trends *prior* to funding and provides flexibility as to the timing of any potential treatment effect. To simplify presentation of our results, we also present results from a modified difference-in-differences model that allows for separate short-, medium-, and long-run effects. We describe the regressions estimated below. Section VII presents the results.

We follow each researcher, i, for t periods, $t = 1, \ldots, T$. Our goal is to understand the effect of treatment (e.g. receiving a grant in our sample) on a scientist's choice of research topics in each year t, denoted $Y_{i,t}$. Funded researchers receive a grant in period t_{FY}, so that they are untreated in periods $t = 1, \ldots, t_{FY\text{-}1}$, and treated in periods $t = t_{\mathrm{FY}}, \ldots, T$. FY represents the year of funding, and indicates the treated cohort to which each PI belongs Our control researchers remain untreated during the entire sample period.

Ideally, our goal is to estimate the average treatment effect on the treated (ATT): E[$Y_{i,t}(1)$ - $Y_{i,t}(0)$ |

$D_{i,t} = 1]$, where $Y_{i,t}(1)$ represents person i's outcome in period t if treated, and $Y_{i,t}(0)$ represents the same person's outcome in period t if untreated. However, because we observe each researcher in only one state (treated or untreated), what we observe is

$$E[Y_{PI,t}(1) \mid D_{PI,t} = 1] - E[Y_{C,t}(0) \mid D_{C,t} = 0],$$

where the subscripts indicate whether researcher i is a treated PI (PI) or control (C), and $D_{i,t}$ is a 0/1 indicator of whether researcher i is in a treated group in time t.

Adding and subtracting the unobserved outcome for the treated researcher ($E[Y_{i,t}(0) \mid D_{i,t} = 1]$), the above expression gives us the treatment effect if the last two terms below (e.g. selection bias) equal 0:

$$E[Y_{PI,t}(1) \mid D_{PI,t} = 1] - E[Y_{PI,t}(0) \mid D_{PI,t} = 1] \; + \; E[Y_{C,t}(0) \mid D_{C,t} = 1] - E[Y_{C,t}(0) \mid D_{C,t} = 0].$$

That is, for our matching to estimate the treatment effect, we must assume that the treated and matched controls would have the same outcome if the funded researcher had not been treated. While this is untestable, we can test whether the topic choices of our treated and control researchers evolve similarly prior to treatment. As the previous section shows, our control group represent good matches for the PIs in our sample. Both controls and PIs are on similar research paths prior to funding. Their paths diverge only after the PIs receive funding. Moreover, most observable characteristics are within acceptable limits for sample balance.

To remove any remaining differences between the PIs and controls, we use an event study regression, denoting researchers receiving funding for grant g from funding program p as the treated group. In an event study regression, individual fixed effects control for any remaining time-invariant differences between PIs and their matched controls. This normalization allows us to compare PIs and controls with the small pre-treatment differences removed. To interpret our results as causal, one must assume that the PIs and controls would have continued to evolve on similar paths after the year of funding *if the PI had not received funding*.

As is standard in the event study literature (Miller, 2023), we normalize the year prior to treatment equal to 0. Here, since the dependent variable is *similarity to a specific grant abstract*, the regression must include researcher-by-grant fixed effects, which treats each researcher/award combination as a separate individual unit in a fixed effect regression. Because a single researcher may receive multiple awards, simply using researcher fixed effects is not sufficient. Since we are measuring similarity to the grant, each award appears as a separate observation in the data, and for PIs receiving multiple awards, the same PI may be matched to different controls depending on the topic of each award.

Let $y_{i,g,p,t}$ represent the similarity of researcher i's publications in year t to their grant abstract, g, which was awarded by program p. $\gamma_{i,g}$ represent the individual-by-grant fixed effects described above. Similarly, we include year effects ($\tau_{t,p,FY}$) specific to each cohort funded by program p. Program-cohort-year fixed effects limit comparisons to those who applied for the same funding competition. For instance, in select years, NSF made additional funds available for clean energy across multiple directorates. Proposals were to be submitted to the appropriate NSF disciplinary area. Thus, made in different years from these NSF programs should be subject to different year fixed effects. Moreover, as discussed below, these ensure only appropriate comparisons are made between PIs and matched controls.

Define funding year, $FY_{g,t}$, as a dummy variable equal to 1 in the start year for grant g. $FY_{g,t}$ equals one in the start year of the grant for both the funded scientist and their matched control. Our treatment variable, $Funded_{i,g,}$ is a dummy variable equal to 1 for the scientist receiving funding from grant g. Here "funded" means receiving a grant from a specific program. It does not mean receiving any research funding. It may be the case that non-treated scientists obtain other funding sources.[2] Thus, our program and cohort specific year fixed effects ($t_{t,p,FY}$) control for general trends affecting both funded scientists and their controls in the same funding cohort, and for each matched pair, the interaction of $FY_{g,t}$ and $Funded_{i,p}$ captures the additional effects of funding on the PI. Using macro-level energy R&D and publication data, previous work finds that the effect of public funding on publication counts may persist for six years or more(Popp, 2016). Our data allow for up to 12 lagged effects (β_{t-s}^{post}) and 12 pre-treatment effects (β_{t-s}^{pre}). However, we bin years 9 and higher into a single coefficient for ease of presentation, and because individual year

[2] In principle, one could control for the size of the grant. Myers (2020) does this in his paper on the elasticity of science. However, as Myers' pool of researchers is all scientists who apply for NIH funding, he knows the funding amounts (or funding applied for) for all scientists in his sample. We do not.

estimates are imprecise at the ends of the sample, when not every researcher is actively publishing. Because they often have large error bars, these binned coefficients are omitted from our figures. The pre-treatment effects (β_{t-s}^{pre}) test for differences in the publication trends of treated and control scientists. Finally, we include controls for time-varying researcher characteristics, to allow for the possibility that the paired researchers are not perfect matches. These controls include a set of dummy variables for institution quality,[3] the average number of citations received per article in the previous three years, and the number of unique co-authors in the previous three years. This gives us the following equation to estimate:

$$y_{i,g,p,t} = \alpha_0 + \sum_{s=1}^{9+} \beta_{t+s}^{pre} FY_{g,t+s} Funded_{i,g} + \sum_{s=0}^{9+} \beta_{t-s}^{post} FY_{g,t-s} Funded_{i,g}$$

$$+\boldsymbol{\phi}\mathbf{X_{i,t}} + \tau_{t,p,FY} + \gamma_{i,g} + \epsilon_{i,t} \quad (1)$$

To allow for differences in the scale of similarity between grant abstracts and publications across different awards, we cluster standard errors by grant *g*.

The event study in equation (1) is our main specification, as it allows us to assess whether PIs and their matched controls were on similar trends prior to funding and provides flexibility as to the timing of any potential treatment effect. To simplify presentation of our results, we also present results from a modified difference-in-differences model that allows for separate short-, medium-, and long-run effects.

$$y_{i,g,p,t} = \alpha_0 + \beta^{SR}\left(Funded_{i,g} * Post_{i,g}^{SR}\right) + \beta^{MR}\left(Funded_{i,g} * Post_{i,g}^{MR}\right) + \beta^{LR}\left(Funded_{i,g} * Post_{i,g}^{L:R}\right)$$

$$+\boldsymbol{\phi}\mathbf{X_{i,t}} + \tau_{t,p,FY} + \gamma_{i,g} + \epsilon_{i,t} \quad (2)$$

Here, the various *Post* variables are 0/1 indicators for whether year *t* is 0-2 (SR), 3-5 (MR, or 6+ (LR) years past the start of funding for grant *g*.

As noted above program and cohort specific year fixed effects are important to ensure that each regression only makes valid comparisons across pairs of researchers. As noted by recent literature on two-

---

[3] While we match researchers on the Carnegie classification of their research institution, here we include additional information on quality using the 2015 Times Higher Education ranking (from https://www.timeshighereducation.com/world-university-rankings) of each researcher's university. We include a set of dummy variables for the top 15 universities, ranking 16-30, 31-50, 51-71, or having no ranking. While we do not match on these rankings to avoid larger reductions in sample size, these rankings allow both an additional control for institution quality and reflect changes in a researcher's home institution over time.

way fixed effects models (Callaway & Sant'Anna, 2021; Sun & Abraham, 2021), estimates of traditional two-way fixed effects models are potentially contaminated by invalid comparisons across pairs of observations. Traditionally, the concern is over early-treated observations being compared to late-treated observations. Here, however, we have a second concern. Our dependent variable measures similarity to a specific grant abstract. Thus, ensuring that the research of our controls is compared to the same document is important. Two researchers working on *identical* sets of topics could nonetheless have different similarity measures depending on which document those topics are compared to. A simple year fixed effect would compare a treated researcher to *all* controls, even those whose dependent variable is similarity to a project on a very different topic. For example, one program may support research relevant to solar energy, whereas a second program may relate to biology, which is fundamental for work on biofuels. In addition, compositional changes in comparison groups can occur in traditional two-way fixed effects event model (Callaway & Sant'Anna, 2021). Callaway and Sant'Anna (2021) address this by "balancing" groups with respect to event time when aggregating average treatment on the treated effects (ATTs). In our specification, cohort-specific effects achieve a similar goal, while still ensuring that sets of PIs are only compared to their matched controls.[4]

One additional complication to a causal interpretation is that the usual assumption of stable unit treatment value assumption (SUTVA) may be violated. SUTVA requires that researcher *i*' outcome does not depend on the treatment status of other researchers. In our case, if a matched control applied for the same funding as the PI and didn't get it, the PI receiving the award instead may cause the control to take up an alternative research topic for which they can get funding. That is, it may be the lack of funding available to the control, rather than the funding received by the PI, influencing our observed effect. Nonetheless, given potential threats to energy funding, understanding how the choice of topics between

[4] To validate that cohort-specific year fixed effects ensure comparisons are made to the correct controls, note that in a regression replacing individual fixed effects with grant fixed effects and omitting any controls, a regression using cohort-specific year fixed effects exactly replicates paired t-test between the PIs and controls in Figure 6. The same is not true if year fixed effects are allowed to vary across cohorts. Moreover, a simple F-test rejects the null hypothesis that the year fixed effects are equal across cohorts.

funded and non-funded researchers evolve is important whether any observed changes result from the PI receiving funding or the control not receiving funding.

VII. Results

We begin by comparing the evolution of funded principal investigators' research topics to those of matched control scientists who had similar research profiles prior to the grant start date. We summarize these dynamics using both an event study (Figure 7) and a difference-in-differences specification that distinguishes short-, medium-, and long-run effects (Figure 8). The event study verifies that PIs and their matched controls were on similar trends *prior* to funding and provides flexibility as to the timing of any potential treatment effect. Finally, we compare estimated effects across funding programs to assess how mission orientation and program design shape the extent to which funding redirects researchers toward new topic areas.

Overall, most energy-specific and targeted programs move research closer to the funded topic than general calls for funding (Figures 7 and 8). Notably, these effects persist over time. This persistence suggests that changes are not merely a result of projects directly funded by the grant, but rather help build human capital by keeping researchers in the field. These results are consistent with prior work finding that increases in energy R&D lead to new energy publications even 10 years after funding (Popp, 2016). In contrast, researchers use general funding to move existing research streams forward, rather than move into new topic spaces.

For ARPA-E and NSF Energy-specific programs, the effects are two-to-three times as large as general programs in each agency. In NSF, energy programs provide a steady increase in similarity, whereas in ARPA-E the largest gains come in the long run. EERE provides modest gains similar to those of NSF energy programs, but with less precision. Effects are only significant at the ten percent level, and only in the short-run (p-value = 0.089) and long-run (p-value = 0.10).

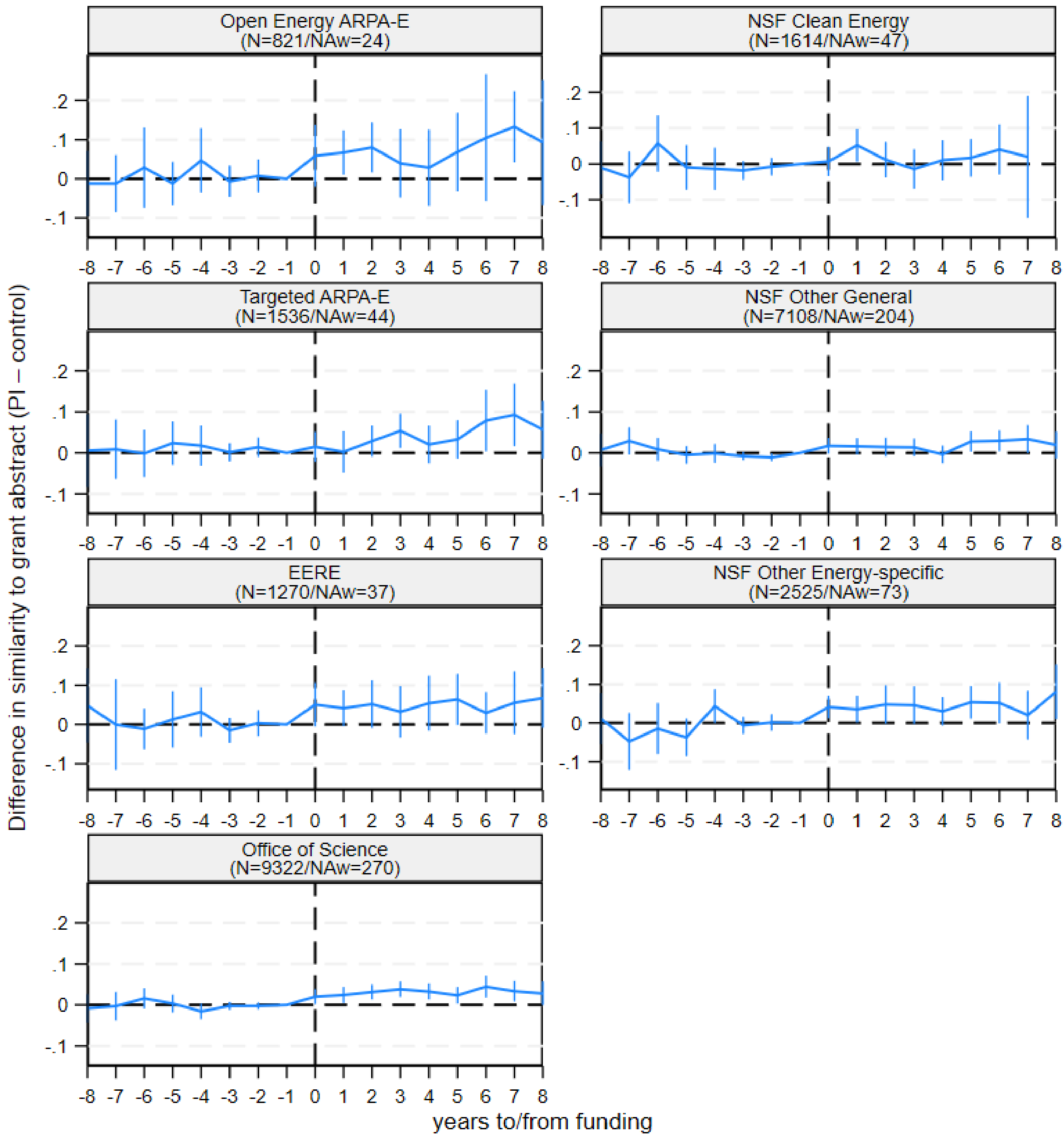


Figure 7 | Event study results. Results of event study regressions comparing the similarity of a PI's publications to their grant abstract, relative to their matched control. N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award.

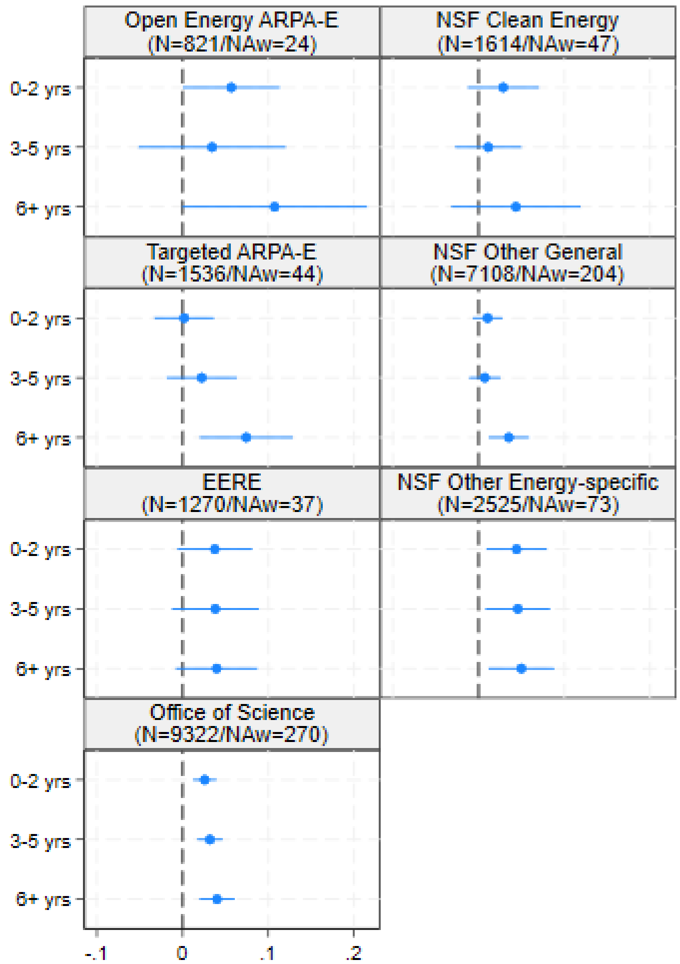


Figure 8 | Effect of Funding on Similarity. Results of difference-in difference regressions comparing the similarity of a PI's publications to their grant abstract, relative to their matched control. N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award.

To interpret the magnitude of our estimates, note that the average PI in our sample has roughly 10 publications per year. Thus, replacing one paper completely unrelated to the funded topic (e.g. similarity = 0) with a paper perfectly related (e.g. similarity = 1) would increase similarity to the funded research by 0.1. Moving similarity by 0.05 per year is the equivalent of replacing, from that set of 10 papers, a paper half-related to the funded topic with one that is fully related. Using this thought experiment, both EERE and NSF Energy-specific programs both provide consistent effects over time, changing research topics by about 0.4 papers per year.

ARPA-E produces the largest changes in research topics with long-run effects between 0.7 and 1 full new paper per year. While the estimates are less precise due to smaller sample sizes, they are statistically significant. Open ARPA-E programs produce faster movement towards the topics than targeted ARPA-E, with magnitudes approaching 0.077 after just two years after funding (Figure 7). By giving PIs more flexibility on what to do within energy research, open energy-specific calls may facilitate new researchers moving into the energy space by reducing the cost of pivoting (Hill et al., 2025). While the long run estimates are less precise due to a smaller sample, Open ARPA-E awards continue to generate a new paper per year related to the funded topic even 6-8 years after funding. ARPA-E uses Open solicitations to avoid missing potential innovative opportunities outside of Focused FOAs. These Open solicitations appear to facilitate ARPA-E's goal of attracting new scientists to energy research.

Within NSF, one interesting observation is the small effect of Clean Energy Awards. In select years, NSF made additional funds available for clean energy across multiple directorates and made the availability of these funds known through a series of Dear Colleague Letters (Appendix A). Proposals were submitted to the appropriate NSF disciplinary area. Other than a slightly larger short-run effect, we see little difference between the effect of these awards and other awards made by NSF's general programs in our sample. Simply making additional funds available for clean energy research is not sufficient to attract new researchers. In contrast, NSF's energy-specific awards solicited by specialized NSF programs, such as Sustainable Energy Pathways, did result in greater movement toward the topics funded.

That our observed effects are both larger for energy-specific programs and persist over time shows that the funding itself matters. Government funding calls are not simply acting as a signal to the broader research community, as our control group does not find other resources to enable them to perform similar research. In contrast, if funding announcements acted as a signal of growing demand for a topic, rejected researchers could attempt to fund their proposals through other means. Such a result would bias our results downward, but would be most notable for targeted programs designed to influence the topics of applications received. Because it would take additional time for failed projects to find funding elsewhere, such a scenario would suggest that funding leads to an immediate movement towards energy research, with the effect disappearing over time as failed applications are funded elsewhere and completed at later dates. That is the opposite of what we observe.

VIII. Mechanisms and Impact

A. Potential Mechanisms

One possible concern is that our finding that funded researchers moving in the direction of their grant is purely mechanical – that PI's subsequent publications are similar to their funded grant abstract by construction. Appendix E provides further analysis removing papers directly acknowledging grant funding. For general award programs, the short-run effect of funding on research direction goes to zero when removing papers acknowledging the grant (Appendix Figure E2) before returning to a long-run upward trend, suggesting researchers use general awards to continue research on topics for which they have previous experience.

For both EERE and NSF energy-specific awards, publications of PIs are more similar to the grant abstract even *in the year of funding* (Appendix Figure E2, EERE: β=0.0484, p-value 0.095; NSF β=0.047, p-value 0.017). Along with finding less change when removing acknowledgement for these programs, these programs appear to fund work consistent with the overall research portfolio of PIs. In contrast, for ARPA-E we observe less movement towards the grant topic when removing papers acknowledging the grant, but not no movement. Moreover, the persistent long-run effects of ARPA-E suggest that bringing scientists into the energy space develops human capital. In sum, general awards seem to mainly help research already underway, but ARPA-E funding tends to shift researchers into new territory, which shows up even in papers that never acknowledge the grant. Moreover, the grants appear to create durable expertise rather than being one project output, which is consistent with the agency's mission.

Next, we look for differences in results based on researcher characteristics. Funding is more likely to change the direction of older PIs than younger PIs (Appendix Figure F1). This difference is particularly notable for NSF general programs, which has no effect on the research direction of younger PIs. Similarly, ARPA-E has no effect on the research direction of young researchers but does lead to shifts in research direction from older PIs. That this effect persists over time provides further evidence that ARPA-E was able to attract established researchers into the clean energy space. This result differs from recent work (Hill et al., 2025) finding a higher pivot penalty for older researchers. For ARPA-E, it may be that younger generations are already working on the emerging problems that ARPA-E identified, whereas more senior scholars have spent decades working on problems that were emerging when they started their career.

Prior research topics may also influence the effect of funding. As would be expected, the effect of funding on similarity to the grant is larger for those whose prior research was dissimilar to the funded topic. For most programs, funding has no effect on the research direction of PIs whose prior work was similar to their grant (Appendix Figure F2). Similarly, grants are more likely to influence research topics of those PIs with the highest research breadth, as measured by entropy (Appendix Figure F3).

Finally, in Appendix G we demonstrate the robustness of our results to key assumptions, including different matching rules, different year fixed effects and different estimation methods.

B. Do New Researchers Produce High-Impact Research?

While ARPA-E succeeded at attracting new researchers to clean energy, are these programs also effective at selecting PIs who produce high-impact work in these new areas? Using articles acknowledging funding from the grants in our sample, Figure 9 presents the marginal effects from a regression showing how a PI's similarity to their grant abstract affects the number of citations received within five years of publication. For this analysis, the unit of analysis is an academic article. The dependent variable, $cites_{a,g,t}$ is the number of citations received by article a, associated with grant g, within five years of the year of publication, t. Allowing for non-linear effects, we include both a linear and squared term for the similarity of a PI's prior research to the grant abstract ($PIsim_g$). Appendix Table G1 provides justification for a non-linear model. To account for author productivity, we control for each PI's number of publications &

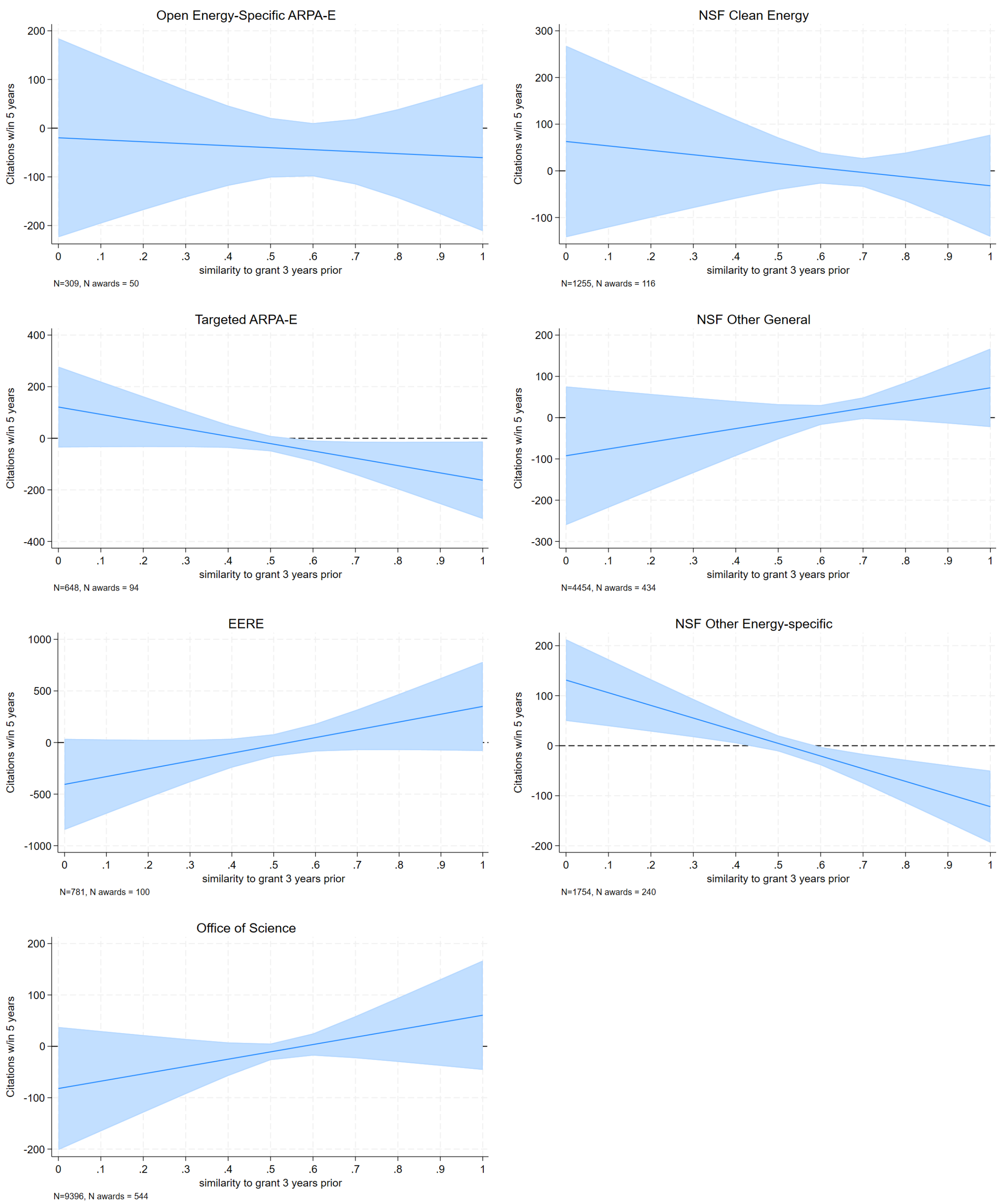


Figure 9 | Marginal effect of similarity on citations. Graphs show marginal effect of similarity on citations. All models control for # of publications & citations in prior 3 years, age at grant, and publication year fixed effects. Standard errors clustered by award number. Shaded region shows 95% confidence intervals.

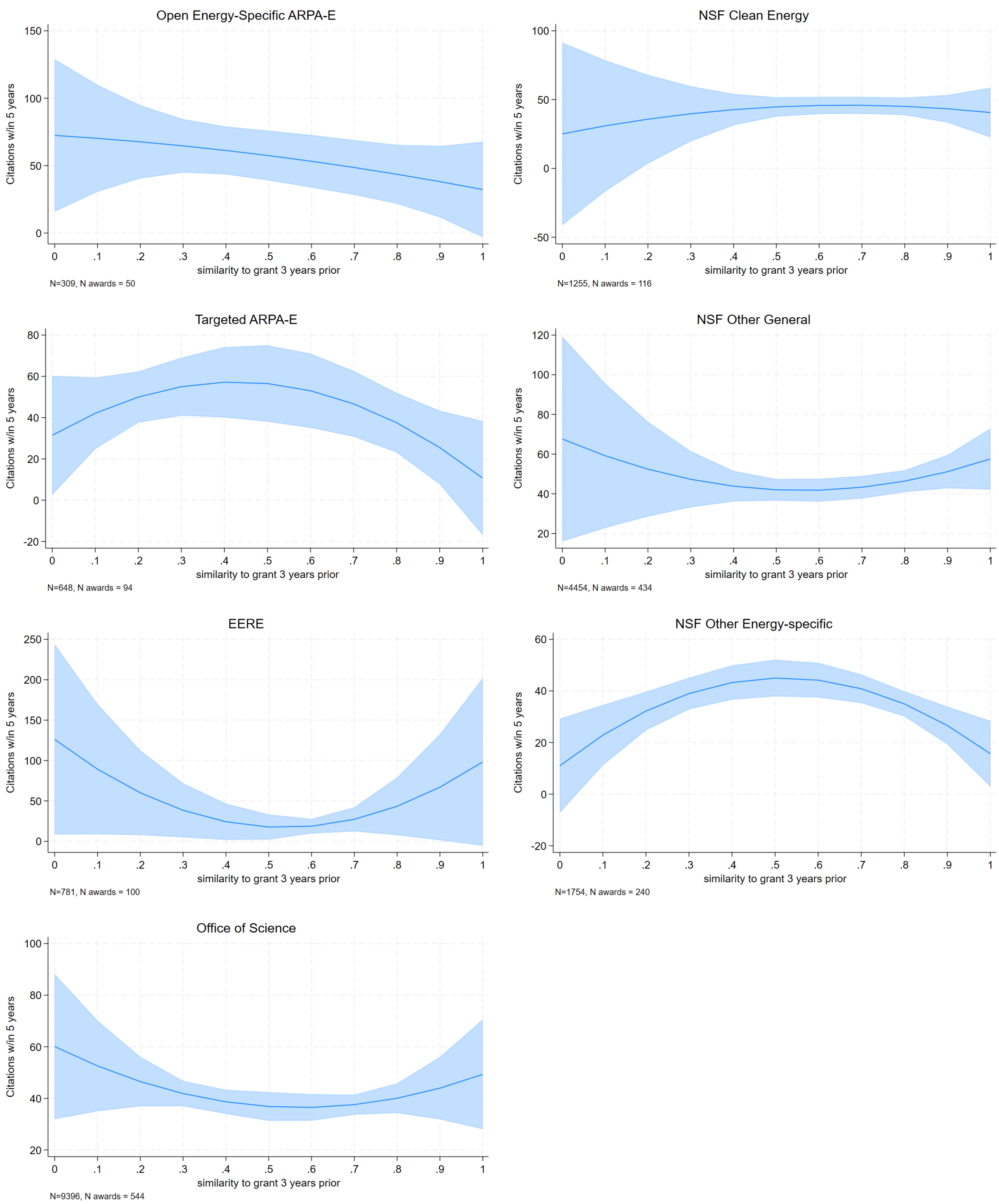


Figure 10 | Effect of similarity on predicted citations.: Graphs show predicted citations. All models control for # of publications & citations in prior 3 years, age at grant, and publication year fixed effects. Standard errors clustered by award number. Shaded region shows 95% confidence intervals.

citations per article in the three years prior to funding and the author's experience at the time of the grant. These are represented by the matrix $\mathbf{X_g}$ below. The regression also includes year fixed effects (η_t). Our estimating equation is:

$$cites_{a,g,t} = \alpha + \beta_1 PIsim_g + \beta_2 PIsim_g^2 + \boldsymbol{\gamma}\mathbf{X_g} + \eta_t + \epsilon_{a,g,t} \qquad (3)$$

We run separate estimations for each of our 7 program groups. In each, standard errors are clustered by grant. Appendix G shows that our results are robust to alternative measures of article quality, including citations counts normalized by field and papers in the top 5 percent of citations for their field.

For EERE, NSF Clean Energy, and NSF general programs, similarity has little impact. While papers published by PIs more distant from their grant receive fewer citations, the effect is never statistically significant at a 5% confidence interval. However, for targeted ARPA-E and NSF energy-specific programs, articles published by PIs more similar to the grant receive fewer citations. For both programs, predicted citations are highest for authors with similarity scores near 0.5 (Figure 10). While attracting new researchers is not always advantageous, it appears most advantageous in programs explicitly seeking new researchers in the clean energy space.

IX. Discussion

Overall, our results show that even if research directions are incremental and path dependent (Hill et al., 2025; Jia et al., 2017), they are not immutable. Funders can craft calls that prioritize problems aligned with their missions to compel researchers to shift their focus. We find strong evidence that grants influence the topical directions pursued by energy researchers funded by DOE and NSF grants – with effects as strong as replacing a paper unrelated to the topic of the grant by one that is a 77% match, just two years after the onset of the grant in the case of Open ARPA-E calls. Grant programs seeking to fund risky but promising ideas yield the largest shifts. Researchers appear keen to work on problems targeted by funding agencies and develop a lasting work program in these areas. The largest shifts occur in the Open ARPA-E program – suggesting that researchers shift direction more easily when given space to propose unconventional ideas for solving the problems identified by funding agencies. Targeted ARPA-E awards produce similar movement, but only in the long-run. Together, these results point to trade-offs that inform the choice of grant designs. If funders aim to attract researchers new to narrowly defined topics in targeted FOAs, it may take longer to produce outputs closely related to the grant than if they give researchers the freedom to propose new solutions to agency-defined problems.

Conventional grants that give researchers greater latitude to choose the problems they focus on yield the smallest shifts, even when additional funds are available for clean energy research (e.g., NSF Clean Energy grants). This is consistent with previous literature finding that the adjudication process of conventional grant programs places greater weight on demonstrating project feasibility and the competence (i.e., prior experience) of the research team (Luukkonen, 2012). In contrast, ARPA-E program managers have discretion to deviate from the recommendations of peer reviewers. Program directors give preference to proposals with greater disagreement among reviewers and that appear more creative (Goldstein & Kearney, 2024). Notably, PIs whose prior work is less similar to the grant produce more highly cited work only when attracting such researchers is a program goal.

Our findings also contribute empirical insights to broader debates about the role of science funders: supporting scientists in advancing the frontiers of science without being constrained by the immediate applicability of their work, or encouraging scientists to focus on 'mission' research with direct applicability for solving pressing societal challenges (Mazzucato, 2018; Rodrik, 2014)? We show that calls soliciting ideas on targeted problems move grantee's research portfolio more than conventional grants, with implications for how funders may want to bundle different types of funding modalities to achieve these different goals. While we showcase examples on energy research, our findings have broader generalizability to high-risk/high-reward funding modalities in any scientific domain. For example, the Advanced Research Projects Agency for Health (ARPA-H) was launched in 2022 to fund breakthrough health research. The United Kingdom launched its Advanced Research Innovation Agency (ARIA) in 2023 to fund breakthrough research across a range of topics, using ARPA as a model. While the specific 'missions' defined by these programs may differ from ARPA-E, the ultimate science policy goal of funding transformative research to find solutions to pressing societal problems is the same.

One important caveat is that our estimates capture changes in research trajectories relative to a matched set of unfunded scientists and reflect how funding reallocates across researchers. In settings with scarce funding, observed divergence may arise both because funded scientists sustain attention to a topic and because unfunded scientists redirect effort elsewhere. Nonetheless, given potential threats to energy research funding, understanding how the choice of topics between funded and non-funded researchers evolve is important whether observed changes result from the PI receiving funding or the control not receiving funding.

Finally, grants not only fund research activities and outputs, they also help build human capital. They develop the expertise of both the current and next generation of scientists in targeted areas of science considered important for society. Our findings, consistent with previous macro-level studies (Popp, 2016) show that grants have a lasting impact - of several years to a decade – on grantees' research programs. Such persistence over time may also reflect follow-on funding, durable research teams, or institutional repositioning enabled by the initial award. In the current US federal policy context, where research advancing energy transitions is at risk, our work offers several important insights about the long-term effects of funding cuts. First, while we observe differences in the research trajectories of funded PIs and their matched controls, our descriptive data (e.g. Figure 6) suggests that these differences are at least in part due to the research of the controls becoming less similar to the funded topic. What is important is that, whatever the mechanism, the funding programs in our sample are not simply funding research on topics that the PIs would have done anyway. Future funding cuts are thus likely to move even more scientists out of clean energy. Second, any pause or redirection of research funding away from sustainability has long term consequences, as it is a missed opportunity to train talent and deepen expertise. It will set scientists onto paths that are difficult to correct. Finally, our findings also offer a glimmer of hope for restoring clean energy research. The ability of high-risk/high-reward grants to induce scientists to change the course of their research trajectories could be used by future administrations who wish to move researchers back into a focus on sustainability and advancing energy transitions.

References

Achakulvisut, T., Acuna, D. E., Ruangrong, T., & Kording, K. (2016). Science Concierge: A fast content-based recommendation system for scientific publications. *PloS ONE*, *11*(7), 0158423.

Aggarwal, C. C., Hinneburg, A., & Keim, D. A. (2001). On the surprising behavior of distance metrics in high dimensional spaces. *Proceedings of the 8th International Conference on Database Theory*, 420–434.

Azoulay, P., Graff Zivin, J. S., & Manso, G. (2011). Incentives and creativity: Evidence from the academic life sciences. *The RAND Journal of Economics*, *42*(3), 527–554.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

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

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics, Themed Issue: Treatment Effect 1*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Chai, Y., & Li, W. (2019). Towards deep learning interpretability: A topic modeling approach. *Proceedings of the 40th International Conference on Information Systems (ICIS*.

Deerwester, S., Dumais, S., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, *41*(6), 391–407.

Delgado-Quirós, L., & Ortega, J. L. (2024). Completeness Degree of publication metadata in eight free-access scholarly databases. *Quantitative Science Studies*, *5*(1), 31–49.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186.

Furman, J., & Teodoridis, F. (2020). Automation, Research Technology, and Researchers' Trajectories: Evidence from Computer Science and Electrical Engineering. *Organization Science*, *31*(2), 330–

354.

Goldstein, A., & Kearney, M. (2024). *Uncertainty and individual discretion in allocating research funds*. NBER Working Paper #32033. http://www.nber.org/papers/w32033.

Goldstein, A., & Narayanamurti, V. (2018). Simultaneous pursuit of discovery and invention in the US Department of Energy. *Research Policy*, *47*(8), 1505–1512.

Gugler, K., Szücs, F., & Wiedenhofer, T. (2024). Environmental Policies and directed technological change. *Journal of Environmental Economics and Management*, *124*, 102916. https://doi.org/10.1016/j.jeem.2023.102916

Hill, R., Yin, Y., Stein, C., Wang, X., Wang, D., & Jones, B. F. (2025). The pivot penalty in research. *Nature*. https://doi.org/10.1038/s41586-025-09048-1.

Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, *95*, 1168–1177.

Jia, T., Wang, D., & Szymanski, B. K. (2017). Quantifying patterns of research-interest evolution. *Nature Human Behavior*, *1*, 0078.

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

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, *84*(5), 905–949.

Luukkonen, T. (2012). Conservatism and risk-taking in peer review: Emerging ERC practices. *Research Evaluation*, *21*, 48–60. https://doi.org/10.1093/reseval/rvs001.

Mancuso, R., & Broström, A. (2026). Do mission-oriented grant schemes shape the direction of science? *Research Policy*, *55*(1), 105360. https://doi.org/10.1016/j.respol.2025.105360

Mandt, R., Seetharam, K., & Cheng, C. H. M. (2020). Federal R&D funding: The bedrock of national innovation. *MIT Science Policy Review*, *1*, 44–54.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and Corporate Change*, *27*(5), 803–815. https://doi.org/10.1093/icc/dty034.

Meckling, J., Galeazzi, C., Shears, E., Xu, T., & Diaz Anadon, L. (2022). Energy innovation funding and institutions in major economies. *Nature Energy*, *7*, 876–885.

Merton, R. K. (1957). Priorities in scientific discoveries: A chapter in the sociology of science. *American Sociological Review*, *22*, 635–659.

Miller, D. L. (2023). An Introductory Guide to Event Study Models. *Journal of Economic Perspectives*, *37*(2), 203–230. https://doi.org/10.1257/jep.37.2.203

Myers, K. (2020). The Elasticity of Science. *American Economic Journal: Applied Economics*, *12*(4), 103–134.

National Academies of Sciences, Engineering, and Medicine. (2017). *An Assessment of ARPA-E*. The National Academies Press.

Office of Energy Efficiency and Renewable Energy. (n.d.). *2016-2020 Strategic Plan and Implementing Framework* (DOE/EE-2061). Retrieved September 2, 2025, from https://energy.gov/sites/prod/files/2015/12/f27/EERE_Strategic_Plan_12.16.15.pdf

Popp, D. (2016). Economic analysis of scientific publications and implications for energy research and development. *Nature Energy*, *1*(4), 1–8.

Popp, D., Pless, J., Haščič, I., & Johnstone, N. (2022). Innovation and Entrepreneurship in the Energy Sector. In A. Chatterji, J. Lerner, S. Stern, & M. J. Andrews (Eds.), *The Role of Innovation and Entrepreneurship in Economic Growth* (pp. 175–248). University of Chicago Press.

Popp, D., Vona, F., Grégoire-Zawilski, M., & Marin, G. (2024). The Next Wave of Energy Innovation: Which Technologies? Which Skills? *Review of Environmental Economics and Policy*, *18*(1), 45–65.

Rijcken, E., Kaymak, U., Scheepers, F., Mosteiro, P., Zervanou, K., & Spruit, M. (2022). Topic modeling for interpretable text classification from EHRs. *Frontiers in Big Data*, *5*, 846930.

Rodrik, D. (2014). Green industrial policy. *Oxford Review of Economic Policy*, *30*(3), 469–491.

Rosenbloom, J. L., Ginther, D. K., Juhl, T., & Heppert, J. A. (2015). The Effects of Research & Development Funding on Scientific Productivity: Academic Chemistry, 1990-2009. *PLoS ONE*, *10*(9), 0138176.

Rubin, D. B. (2001). Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology*, *2*, 169–188.

Shin, H., Kim, K., & Kogler, D. F. (2022). Scientific collaboration, research funding, and novelty in scientific knowledge. *PLoS One*, *17*(7), 0271678.

Stephan, P. E. (1996). The Economics of Science. *Journal of Economic Literature*, *34*(3), 1199–1235.

Stephan, P. E. (2010). The Economics of Science. Handbook of the Economics of Innovation: Vol.1. In B. Hall & N. Rosenberg (Eds.), *Academic Press/Elsevier* (pp. 217–273).

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, Themed Issue: Treatment Effect 1*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

van den Heuvel, M., & Popp, D. (2023). The role of venture capital and governments in clean energy: Lessons from the first cleantech bubble. *Energy Economics*, *124*, 106877. https://doi.org/10.1016/j.eneco.2023.106877

Verdolini, E., & Galeotti, M. (2011). At Home and Abroad: An Empirical Analysis of Innovation and Diffusion in Energy Technologies. *Journal of Environmental Economics and Management*, *61*, 119–134.

Zhao H., Phung D., Huynh V., Jin Y., Du L., & Buntine W. (2021). *Topic modelling meets deep neural networks: A survey*.

Supplementary Information for "Government Funding and the Direction of Academic Energy Research"

Section A: Detailed Description of Research Programs	A1
Section B: Measuring Similarity and Research Breadth	A7
Section C: Further Description of the Matching Process	A16
Section D: Evaluation of Matching	A20
Section E: Results Removing Papers Acknowledging Grants	A29
Section F: Results by PI Characteristics	A32
Section G: Robustness Checks	A35
References	A50

A. Detailed Description of Research Programs

Our analysis of the effect of funding on research direction compares programs relevant for clean energy at two agencies: the U.S. Department of Energy and the National Science Foundation. We include *general* programs that solicit research on a range of topics, *energy-specific* programs soliciting research on clean energy topics defined by researchers, and *targeted* energy programs where researchers respond to specific funding requirements. Table 1 in the main paper, repeated below, summarizes our categorization of the programs included in our study, which we describe in more detail below.

General	DOE, Office of Science Basic Energy Sciences *NSF General:* • NSF, Chemical Catalysis (except NSF Clean Energy) • NSF, Electronic & Photonic Materials (except NSF Clean Energy)
Open Energy-specific	*ARPA-E Open Energy-Specific* • DOE ARPA-E Open • DOE ARPA-E IDEAS *NSF Open Energy-Specific* • NSF Clean Energy Technology (awards made by NSF Chemical Catalysis or Electronic& Photonic Materials) • NSF Energy and Environment • NSF Energy for Sustainability • NSF Sustainable Energy Pathways
Targeted Energy-specific	DOE ARPA-E Focus DOE EERE: Bioenergy Technologies* DOE EERE: Solar Energy Technologies DOE EERE: Wind Energy Technologies DOE EERE: Water Power Technologies

**: One EERE solicitation in 2014 included an Open call for biofuels proposals. We code awards made in response to that FOA as Open Energy-specific, but do not present separate results for those awards given the small number.

Table 1 | Categorization of research programs

We include three programs at the DOE spanning this range of funding options. The DOE's Advanced Research Projects Agency-Energy (ARPA-E) began operating in 2009. ARPA-E aims to identify and fund high-risk, potentially high return research in technologies that will "reduce imports of fossil fuels, reduce energy-related emissions, improve energy efficiency ... and ensure that the United States maintains a technological lead in the development and deployment of advanced energy technologies" (National Academies of Science, 2017, p. 25). Among ARPA-E's goals is creating new communities of researchers, including those "not traditionally... involved in the topic area" (National Academies of Science, 2017, p. 122). Thus, changing the direction of research is an explicit goal of the ARPA-E program.

ARPA-E solicits research proposals through two types of Funding Opportunity Announcements (FOA). Focused FOAs target specific technology areas. These FOAs include clear, well-defined project goals for potential applicants. For example, the FOA for ARPA-E's Green Electricity Network Integration program (DE-FOA-0000473) includes four pages of technical performance targets that successful applications must demonstrate an ability to meet. Our data include awards made by 34 focused programs with start years between 2010 and 2016. *Open* FOAs solicit proposals for potentially disruptive technologies across a wide range of energy applications. Their purpose is to ensure that ARPA-E does not miss potential innovative opportunities in areas outside of the focused FOAs. Open FOAs account for about one-third of ARPA-E's funding support, which were awarded during three open solicitations in our sample (2009, 2012, & 2015). Finally, in 2013 and 2016 ARPA-E solicited proposals under the IDEAS program, short for Innovative Development in Energy-Related Applied Science. IDEAS provides an opportunity for the rapid support of early-stage applied research to explore pioneering new concepts with the potential for transformational and disruptive changes in energy technology. The awards, limited to one year in duration and $500,000 in funding, are intended for exploratory or proof-of-concept research.

Because ARPA-E was designed to bridge the gap between basic and applied research at DOE (Goldstein and Narayanamurti 2018), we include additional DOE programs on either side of this spectrum. DOE's Office of Science supports basic research (Goldstein and Narayanamurti 2018). We focus on research funded by the Basic Energy Sciences program.[9] The Basic Energy Sciences program "supports basic scientific research to provide foundations to lay the foundations for new energy technologies and to

[9] Note that publications are an important output of both ARPA-E and Office of Science research. Goldstein and Narayanamurti (2018) show that ARPA-E researchers are more likely to publish than researchers from either EERE or Office of Science. While EERE research is more applied, and often results in technical reports and patents, the EERE projects studied by Goldstein and Narayanamurti are more likely to have a publication than a patent, and we can compare the similarity of the proposed research to a scientist's previous publications even if the EERE grant in question does not produce any publications itself.

advance DOE missions in energy, environment, and national security."[10] It supports both projects with immediate applications to energy (such as DE-SC0005418, "Computer Simulation of Proton Transport in Fuel Cell Membranes" or DE-SC0005392, "Single-Molecule Spectroelectrochemistry of Interfacial Charge Transfer Dynamics In Hybrid Organic Solar Cell"), as well as more fundamental research (such as DE-SC0007064, "Electric-Field Enhanced Kinetics in Oxide Ceramics: Pore Migration, Sintering and Grain Growth" or DE-SC-0007091, "Computational time-resolved and resonant x-ray scattering of strongly correlated materials"). Illustrating that these programs are open to a range of ideas, in recent years the Office of Science has issued a single FOA entitled "Continuation of Solicitation for the Office of Science Financial Assistance Program," which provides brief one or two paragraph descriptions of each program. DOE's Office of Energy Efficiency and Renewable Energy (EERE) is the applied research program supporting clean energy projects most comparable to those funded by ARPA-E. Like ARPA-E, EERE uses targeted FOAs to solicit research proposals that specify detailed technological improvements to be achieved. Its goal is to address gaps in technology development where the private sector will not invest (Office of Energy Efficiency and Renewable Energy, 2015).

NSF standard grants include multiple directorates overseeing programs supporting research in a broad range of science and engineering fields through open calls for proposals. Given the large number of programs in NSF, we focus on two programs commonly used as funding sources by researchers also funded by DOE, so that we are comparing NSF programs that our DOE funded researchers are likely to consider. These are the Chemical Catalysis and Electronic & Photonic Materials Programs. Both programs attract researchers working on energy, but they also fund related research on other topics. This provides an opportunity to see if broad programs can attract scientists interested in switching topics. Furthermore, in select years, NSF made additional funds available for clean energy across multiple directorates. On the next page, we provide a sample Dear Colleague Letter (NSF 14-115) informing researchers of funding available for clean energy technology in multiple directorates. Proposals were to be submitted to the appropriate NSF

[10] https://www.energy.gov/science/bes/basic-energy-sciences, last accessed February 15, 2025.

disciplinary area. These are identified in NSF's funding database with reference code 8396. Using this code, we separately identify such awards made to these two directorates.

While most NSF programs use open call for proposals, a few programs specifically solicit research pertaining to clean energy. Unlike targeted DOE programs, these NSF programs solicit proposals for a range of clean energy technologies, and are thus classified as Open Energy-specific programs. As the audience of potential grantees at NSF is likely larger than DOE, understanding the impact of such programs at NSF is important. First, we include awards related to Energy and the Environment made through the Industry–University Cooperative Research Centers (IUCRC) program. The IUCRC program aims to accelerate the impact of basic research through partnerships between industry, academics, and government. They support research in several focus areas, one of which is Energy and the Environment.[11] Second, we include awards through the Energy for Sustainability program. This program supports fundamental engineering research to enable sustainable production of electricity and fuels.

Third, NSF offers various cross-cutting solicitations in which multiple programs participate. Relevant for this proposal is 2011's "Sustainable Energy Pathways" program. Unlike the aforementioned energy programs in NSF, which are in a single program, Sustainable Energy Pathways involved five different NSF directorates and required interdisciplinary teams of researchers. It is an open solicitation, explicitly mentioning interest in a "broad scope of sustainable energy interest areas." The solicitation does not provide specific technological goals

[11] https://iucrc.nsf.gov/, last accessed February 7, 2025.

Sample Dear Colleague Letter (https://www.nsf.gov/funding/opportunities/dcl-fy-2015-clean-energy-technologies-funding-opportunities/nsf14-115, accessed October 9, 2025)

FY 2015 Clean Energy Technologies Funding Opportunities

August 27, 2014

Dear Colleagues:

It is critical to provide sustainable and economical energy systems on a scale sufficient to power all of society's needs. The development of clean energy technologies is an important step in that direction as it addresses the interrelated challenges of producing safe and responsible energy sources while reducing our dependence on foreign oil and minimizing the impact on the environment.

All of the Divisions in the following Directorates are participating in clean energy technology research and education through ongoing funding opportunities: Biological Sciences (BIO), Engineering (ENG), and Mathematical and Physical Sciences (MPS).

For BIO: fundamental research topics of interest in clean energy technology include, *but are not limited to*: systems and synthetic biology to streamline and scale the metabolic and energetic potential of living organisms such as microbes, fungi, algae and plants to produce non-petroleum based sources of important chemicals/materials, feedstocks and fuels. Investigations to assess the impact of fuel and/or bio-renewable chemical production on genome stability, fitness, and phenotype of the production organisms are of interest, as are studies to assess the potential environmental impacts of these technologies.

For ENG and MPS: examples of fundamental research topics of interest in clean energy technologies include, *but are not limited to*: hydrogen generation and storage; biological, chemical, and catalytic conversion of renewable carbon sources (such as biomass, methane, and carbon dioxide); the development of methods and materials that increase energy efficiency, such as the replacement of stoichiometric with catalytic processes; energy storage, transmission, or distribution (e.g. smart grid); power-electronic and energy-conversion devices; fuel cells; solar energy capture and conversion (including biological and bio-inspired processes for the conversion of sunlight to fuels, electricity, or thermal energy); wind/wave/tidal energy; nuclear energy; studies of energy efficiency and use; and carbon dioxide sequestration and storage.

Within these general guidelines, the Directorates encourage the submission of proposals in the areas of clean energy research. Proposals should be submitted to the NSF program appropriate to the disciplinary area of the proposed research in accordance with the submission window and conditions of that program.

Proposals are welcome from either single or multiple investigators. Interdisciplinary proposals that involve principal investigators traditionally supported by different participating divisions are encouraged. Please follow the guidelines and program descriptions located on the NSF website.

Proposals may be submitted in combination with other solicitations. For example, if there are strong collaborations with industry, the Grant Opportunities for Academic Liaison with Industry (GOALI) solicitation can be used in conjunction with this effort. Similarly, proposals may be submitted in combination with the Faculty Early Career Development (CAREER) or the Research in Undergraduate Institutions (RUI) solicitation. Other NSF funding mechanisms such as Early Concept Grants for Exploratory Research (EAGER) and Integrated NSF Support Promoting Interdisciplinary Research and

Education (INSPIRE) may also be appropriate. Principal investigators are urged to consult with the cognizant program officers for additional guidance.

To see examples of awards made in this area visit the NSF Award Abstracts Database and perform a key word search. Alternatively, please visit the webpages of the disciplinary programs of interest in the participating divisions.

We are excited by the opportunities in the clean energy technologies area and encourage our communities to contribute to our sustainable and secure energy future.

Fleming Crim
Assistant Director
Directorate for Mathematical & Physical Sciences

Pramod Khargoneker
Assistant Director
Directorate for Engineering

John Wingfield
Assistant Director
Directorate for Biological Sciences

B. Measuring Similarity and Research Breadth

We can analyze whether funding affects research patterns by detecting changes between sets of documents published before, during, and after a grant's funding period. These changes can be detected using similarity measures. For example, a research change occurs if research becomes more similar to the grant abstract after funding is received. The procedure we describe is agnostic to the sets considered, and therefore enables both comparisons within a scientist or even across a field. We use these similarity measures to assess how the composition of each scientist's research topics changes over time.

There are multiple ways to compute similarities. In this project, we use the text of the grant's abstract and the abstract of each researcher's publications as the primary data to estimate similarities. Specifically, we rely on TF-IDF vectors combined with Latent Dirichlet Allocation (LDA) topic modelling, because this pairing offers transparent representations that can be inspected and validated against researchers' known expertise (Blei et al., 2003; Griffiths & Steyvers, 2004; Achakulvisut et al., 2016). For example, we can perform qualitative validity checks—e.g., confirming that an ARPA-E awardee loads heavily on relevant clean energy research topics—before concluding. In practice, TF-IDF + LDA runs at least an order of magnitude faster than transformer-based (i.e., deep learning) pipelines. It has no hard text limit, making it feasible for millions of full-text records without GPU clusters (Devlin et al., 2019; Beltagy et al., 2019). Recent benchmarks show that, for long documents where topical overlap matters most, "bag of words"-based similarity, such as TF-IDF + LDA, matches or surpasses Transformer-based models (i.e., deep learning) while using a fraction of the compute budget (Rau et al., 2024).

Measuring similarities using NLP involves two steps: (1) transforming the sets of documents into a vector representation, and (2) using an appropriate distance function for estimating their similarities. Ideally, the vector representation should aim to capture the semantics of the documents using a low dimensional space. Some representations such as vectors holding the counts of unigrams or *n*-grams are very high dimensional, causing several difficulties (Aggarwal, Hinneburg, & Keim, 2001). For this reason, dimensionality reduction techniques are performed first (Manning et al., 2008). As for the choice of the distance function, the cosine similarity is preferred because it allows comparison of documents of different lengths. Using dimensionality reduction and cosine similarity, documents can be compared effectively.

The dimensionality reduction used in this project is as follows. First, a matrix X containing the term frequency–inverse document frequency (TF-IDF) of all documents is created. This is, for the document i and term j, the matrix is constructed as

$$X = \left(\mathrm{tf}_{ij} \left(\log \left((1 + n)\left(1 + \mathrm{df}_j\right)^{-1} \right) \right) \right)_{ij} \tag{1}$$

where tf_{ij} is the frequency of term j in document i, n is the number of documents, and df_j is the number of documents with term j. Assuming that there are m terms, the matrix X be an n by m highly-sparse matrix, where m typically ranges in the hundreds of thousands.

After this step, a non-negative matrix factorization (NMF) (Murphy 2012) is performed on matrix X. NMF tends to capture a more semantically meaningful representation of a document compared to using term frequencies or keywords. In order to drastically reduce the dimensions, the matrix X is represented by a set of *topics* T which are in turn mapped back into the full *vocabulary* through V. The matrix V therefore captures correlations between terms. This cannot be achieved by methods that rely on simple term counting, such as the representation used by the PubMed related articles features (*pmra*). We perform the factorization as follows

$$X_{n\times m} = T_{n\times k} V_{k\times m}, \tag{2}$$

where $k \ll m$, making T a "tall-and-skinny" matrix, and V a "wide-and-short" matrix. The matrices T and V ($T \geq 0$, $V \geq 0$) are estimated by solving the following optimization problem: $\min_{T,V} \|X - TV\|_2$. By representing the documents using T instead of the full X, we reduce the dimensionality of the documents from m to k dimensions. The dimension k is typically found by cross validation (Manning et al., 2008).

The representation provided by NMF is closely related to Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA). It makes a more correct assumption (i.e., non-negativity of X) compared to LSA and it is much faster and accurate to compute than LDA (Lancichinetti et al., 2015). Because the matrix T can be casted as a probabilistic representation of all topics present on a document, it can represent the topics of a corpus by simply averaging the topics of each document in the corpus as follows:

$$c = \frac{1}{L}\sum t_z \quad (3)$$

where L is the number of documents in the corpus and t_z is the vector representation of the z-th document in the set. Because the matrix T can be cast as a probabilistic representation of all topics present on a document, it can also represent the topics of a corpus by simply averaging the topics of each document.

LDA data sampling and model selection

We need a set document to train our LDA (n parameter in Eq. 2) and also, we need to determine the right number of topics to represent this set of documents (the k parameter in Eq. 2). To this end, we constructed a dataset that includes publications from both funded researchers (principal investigators of past grants) and a selected set of potential researchers who were thematically similar to funded researchers but did not receive funding. Bibliometric sampling techniques, such as bibliographic coupling and co-citation analysis, were used to identify potential researchers (Boyack & Klavans, 2010; Tang et al., 2008). Specifically, we selected researchers whose publications either cited or were cited by publications authored by funded researchers, leveraging citation-based methods for thematic similarity.

To ensure that potential researchers represented realistic grant competitors, we imposed eligibility criteria focusing on researchers who authored at least 10 research publications, had a U.S. affiliation, and had at least one publication between 2009 and 2016, ensuring active researchers during the funding period were included. Following this filtering, we identified 723,011 potential researchers and retrieved their publications alongside those of funded researchers.

The final dataset included 320,454 publications from 1,965 funded researchers, ensuring that the topic model was trained on a broad yet relevant corpus. For LDA training, we also sampled additional publications beyond those authored by funded researchers to create a more balanced dataset. Publications of the potential controls were sampled to approximately match the volume of papers from funded researchers, reducing potential bias in topic representation. This sample helped us to represent a fuller distribution of words and topics. This diversity helped to more precisely represent distances in the matching process later on (see below).

Text Preprocessing. To prepare textual data for LDA modeling, we applied a preprocessing pipeline consisting of several sequential steps. Standard text preprocessing methods, including stopword removal, lemmatization, and part-of-speech filtering (we include nouns, verbs, adjectives and adverbs), were applied to improve topic model quality (Manning et al., 2008). We began by removing punctuation marks such as commas, periods, exclamation marks, and question marks, as they contribute little to topic modeling. All words were then converted to lowercase to ensure consistency. We subsequently generated bi-grams (e.g., "climate change") using Gensim's Phrases model with a minimum count of 5 and threshold of 300 to improve topic coherence.

Applying part-of-speech filtering allowed us to retain only nouns, adjectives, verbs, and adverbs, which are key components for text representation in topic modeling (Manning et al., 2008). Finally, we implemented lemmatization using spaCy to reduce inflected forms to their base forms.

Training and Selecting Topic Models. We trained LDA models with various hyperparameter settings to determine the most coherent topic structure. A grid search was performed across different topic numbers, α (document-topic prior), and η (topic-word prior), evaluating models using coherence scores. The final model was trained using Gensim with 1,000 iterations, optimizing via online variational inference to efficiently scale to our dataset. The resulting model generated topic distributions for all publications and grant abstracts.

The best final parameters were found for $k = 110$ (topics), $\alpha = 0.6$, and $\eta = 0.6$. A qualitative manual review also confirmed that the model was able to effectively represent specialized scientists, scientists outside the field of energy, and scientists who are very broad in their research portfolio. Appendix Table B1 shows the ten most probable terms for the topics included in Figures 2 &3. These figures and the topics that support them provide evidence that our topic model clusters papers into meaningful categories relevant to the range of research supported by the funding programs in our data.

Appendix Table B1: Most probable terms in top topics

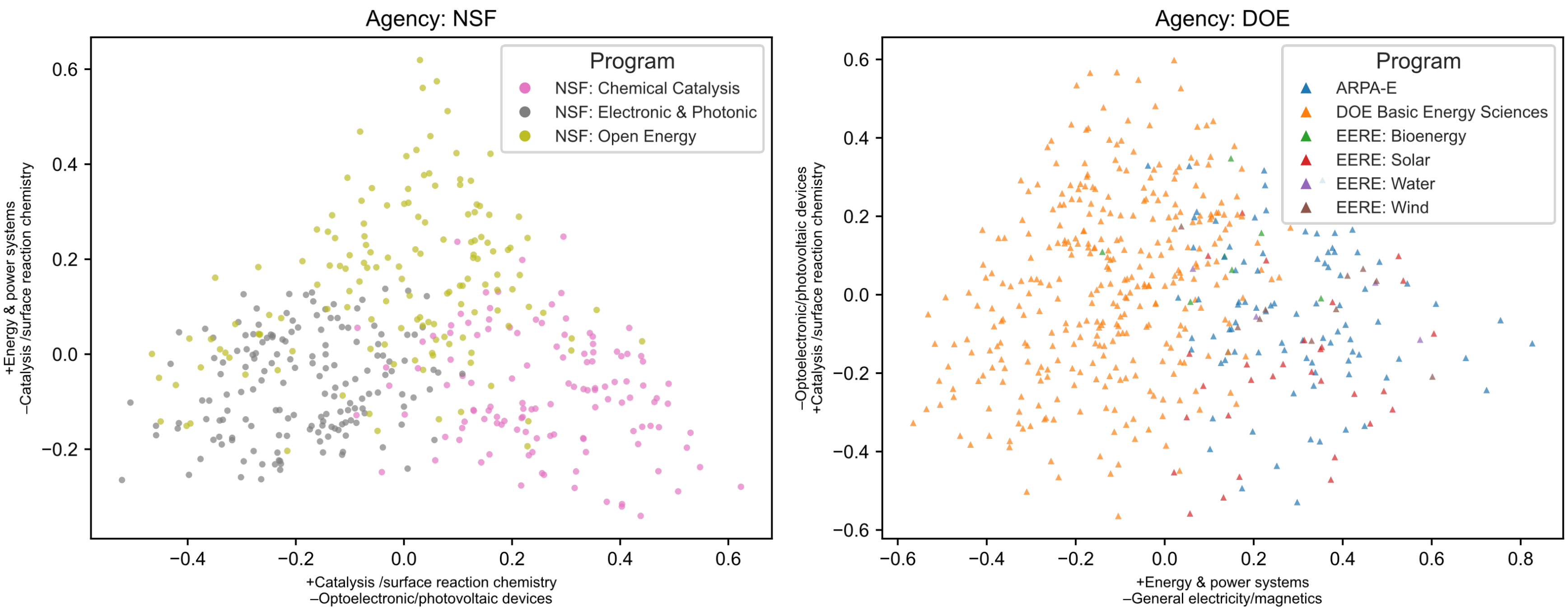
Topic Name	Top 10 Terms
Batteries & energy storage	electrode; ion; battery; electrochemical; high; charge; cell; capacity; cycle; nanowire
Bioprocess engineering	sub; production; sup; high; microbial; bacteria; acid; produce; activity; use
Catalysis /surface reaction chemistry	reaction; complex; catalyst; metal; bond; co; oxidation; site; catalytic; form
Catalytic organic synthesis	compound; ring; product; derivative; reaction; yield; group; substitute; synthesis; catalyze
Control theory	dynamic; theory; function; model; show; force; order; system; motion; wave
Data analysis	network; use; propose; base; image; datum; problem; approach; time; information
Energy & power systems	power; design; energy; control; use; cost; propose; base; system; performance
Environmental impact	area; environmental; impact; food; population; community; use; site; climate; specie
Fuel processing & fluid dynamics	pressure; gas; temperature; co; fuel; heat; high; oil; liquid; use
General electricity/magnetics	state; energy; magnetic; field; spin; structure; transition; temperature; band; density
Generic scientific terms	material; review; provide; new; research; application; recent; development; discuss; property
Genomics	protein; gene; expression; function; cell; enzyme; pathway; identify; mutation; mutant
Materials: thin-films and crystals	film; layer; surface; temperature; growth; thin; substrate; high; structure; use
Optoelectronic/photovoltaic devices	device; high; optical; material; use; demonstrate; light; base; application; laser
Plant physiology/agronomy	plant; soil; growth; high; content; root; increase; seed; yield; tree
Polymer membranes & ions	polymer; membrane; chain; solvent; molecular; concentration; ion; use; aqueous; salt
Simulation/modeling	model; use; result; datum; parameter; simulation; experimental; time; predict; value
Structural biology	bind; protein; structure; site; enzyme; alpha; peptide; residue; fold; beta
Thermal material properties	surface; particle; temperature; use; material; size; high; property; nanoparticle; result
Wind energy/turbines	flow; velocity; wind; speed; wall; jet; energy; turbulence; turbulent; fluid

NOTES: Table shows the top topics by their average appearance across papers funded by each program, as shown in Figures 2 & 3, along with the ten most probable terms in those topics. As in Figures 2 & 3, topics with an applied energy or environmental focus are highlighted in green.

As additional validation, we perform a qualitative analysis of how grants cluster in topic space using a principal component analysis (PCA). We project grant-level topic proportion vectors into two dimensions for NSF and DOE, separately. As relevant topics vary across programs, we use a more detailed breakdown of programs for this analysis. We include separately NSF grants from each of our two general programs (Chemical Catalysis and Electronic & Photonic Materials). Similarly, EERE awards are identified by technology (Bioenergy, Solar, Water, or Wind). Because PCA finds orthogonal directions that capture the largest variance in the data, the plot shows the two most dominant, systematic differences in topic mixtures while reducing the influence of noise. This two-dimensional projection produces a compact and easy to interpret view of the main structure of grants. Each axis is a weighted linear combination of the topics. The X and Y axes contain the highest-loading topics: "+" indicates topics that increase along the positive direction of the axis and "-" along the negative direction. Each point is a grant (Figure B1). Euclidean distances reflect approximate dissimilarity in the topic composition of a grant while clusters indicate grants with similar topical profiles.

For both agencies, the PCA reveals meaningful clusters. Within NSF, the two main programs cluster as they should: Catalysis/surface reaction chemistry is the highest loading topic for the Chemical Catalysis program, and Optoelectronic/photovoltaic devices the highest loading topic for the Electronic and Photonic Materials program (X-axis). Along the Y-axis, awards pertaining to energy and power systems predominately go to the NSF Open Energy programs.

Within the Department of Energy, we see clusters dividing basic and applied energy. Energy and power systems are the highest loading topic for ARPA-E and EERE. Both programs focus on energy applications. In contrast, for DOE's Basic Energy Sciences, which funds basic sciences related to energy, general electricity/magnetics is prominent. Along the y-axis, EERE solar awards cluster in the lower right (Optoelectronic/photovoltaic devices), while bioenergy is in the upper right (Chemical Catalysis program).



Appendix Figure B1 | PCA analysis of grant topics per agency. Each dot represents a grant abstract projected into the first two dimensions of a PCA analysis (left panel NSF, right panel DOE). The X axis represents the most important dimension, and the Y axis represents the second most important. We take the two highest loading topics per dimension and show them as axis labels. A positive sign means that moving to the right increases the present of that topic for the grant, and a negative sign means that moving to the left does so. For NSF, our topic model separately identifies projects supported by the Chemical Catalysis and Electronic & Photonics programs. Along the Y axis, energy grants cluster towards the top. For DOE, the main distinction is between applied energy awards from ARPA-E and EERE (on the right of the X axis) versus basic science research supported by the Office of Science Basic Energy Sciences program (on the left of the X axis).

Measuring Similarity and Breadth of Research. With topic distributions computed for both research papers and grant abstracts, we developed a matching framework to identify researchers whose expertise closely aligned with specific funding opportunities. Each researcher's expertise was characterized by aggregating the topic vectors of their past publications. Since research focus can evolve over time, we constructed separate researcher profiles based on publications from specific time windows before the funding opportunity date.

When comparing authors among themselves or to grant abstracts, we have two sets of documents (i.e., corpora) that we want to compare. We use the cosine similarity between to compare them. More specifically, assume that the corpora of one entity is c_1 and the corporate of the other entity is c_2, we use:

$$\text{similarity}(c_1, c_2) = \frac{1}{2}\left(\frac{c_1 \cdot c_2}{\|c_1\|\|c_2\|} + 1\right). \tag{4}$$

Our similarity measure varies from zero (the two sets are completely different) to one (the two sets are exactly the same).

In addition to these measures of similarity, we also require measures of the breadth of a researcher's portfolio. Among two researchers with comparable similarity measures to a given paper topic, a researcher with additional experience in other areas may be able to combine the knowledge from each field in novel ways. Our breath measure allows us to ask whether researchers working on a wider range of topics contribute more valuable knowledge than other researchers. We first construct a topic distribution of the body of a researcher's work (e.g., set of publications or grants, or both) to construct a topic distribution of it. We can use the normalized vector representation of a set of documents (e.g., using Eq. 3) to estimate the breath of the topics represented in such vector. More specifically, each element of the vector representation s is positive and therefore by normalizing it, we transform the topic assignment into a probability distribution. We then use the entropy of such probability distribution to measure the research breadth. This entropy quantifies how far away the body of work is from a very topically-concentrated distribution, an intuitive proxy for research breadth.

$$\text{entropy} = -\sum_{z=1}^{k} \frac{s_z}{\|s\|} \log\left(\frac{s_z}{\|s\|}\right) \tag{5}$$

C. Further Description of the Matching Process

To establish causal links between funding and the direction of research, we need to know whether the funded scientists pursued research directions considerably different from what they would have done in the absence of the grant. This requires a control group of scientists not funded by these programs to account for other external factors affecting the choices of both funded and non-funded scientists. The universe of all other scientists are not an appropriate control group for our regression, since many would not have been the right candidates to pursue the research funded by these grants in the first place. The best available counterfactual are scientists who were not funded by these research programs, but were on the same trajectory as the sampled researchers in terms of the topical content of their research portfolio, making them just as likely to have pursued the same research directions at the time of the grant. We identify such a control group of authors who are similar to funded scientists in research interest, career stage, and other features *up to the time in which our treated scientists start receiving funding*. We can then compare the research profiles of our funded scientists and matched controls after the start of the funded scientist's grant.

We follow three steps to efficiently identify potential control scientists. First, we obtain publication data on a set of *potential controls* who either cite or are cited by our funded researchers. Second, we further narrow this set of potential controls to *feasible controls* by identifying controls with similar observable characteristics, such as publication counts, research institution quality and age at the time of grant. Third, we use topic modeling to identify *matched controls*, obtaining at most a single match for each funded researcher whose research is both similar overall to the PI ($pairdist_3$) and is a similar distance from the funded research ($WAdist_3$).

<u>*Step 1) Create the universe of potential controls:*</u> First, we develop a set of *potential controls*. As any potential control should work on research similar to our funded scientists, we begin with scientists who either (1) have publications citing articles by our funded researchers or (2) whose publications are cited by our funded researchers. To keep the size of potential controls manageable, we keep only those authors who have at least one affiliation with a US institution during their career, have published more than ten articles, and have at least one publication between 2009 and 2016. We identify 723,011 scientists as potential

controls for our funded scientists.

Step 2) Identify feasible controls with similar observable characteristics to funded scientists: Because processing a large text database to calculate research similarity is computationally expensive, we further narrow the set of *potential* controls using other observable characteristics. These characteristics include characteristics likely to affect a researcher's willingness to change research topics. For both PIs and potential matches, we drop from the data:

1) researchers without a US academic affiliation between 3 years prior to 5 years after the grant
2) researchers with 0 publications in any year three years prior to 5 years after the grant
3) researchers with less than 2 publications per year in the three years prior to 5 years after the grant
4) researchers with less than 3 years of publications prior to the grant
5) researchers with less than 5 years of publications after the grant
6) grants with abstracts of less than or equal to 50 words (similarity to grant analysis only)
7) pairs who are co-authors on a paper acknowledging a grant, or *if no papers acknowledge that grant*, pairs who are co-authors on any paper 0-5 years from the grant start date

As the set of potential controls may also include PIs in our data, we remove any PIs from the set of potential controls.

We then apply the following filters to each PI and remaining potential matches:

- We require that both the PI and match must be at a US academic affiliation. If they have multiple affiliations in their publication record, over 50% must be a US academic institution. Both must be at academic institutions of similar quality in the year prior to the grant, where quality is measured using the Carnegie classification for higher education.
- Defining experience as the number of years since a researcher's first publication, the experience of the PI and potential match at the time of the grant must be +/- 50% of each other.
- Each PI and feasible match must have a similar number of publications in the three years

prior to funding. We keep only potential controls whose 3-year publication count is within +/- 50% of that of their paired PI.

- To ensure similar breadth of research portfolios, the entropy of the PI and potential match in the 3 years prior to the grant must be +/- 25%.

This leaves us with a set of 38,612 feasible matches.

Step 3) Apply results of topic modeling to identify matched controls: For the PIs and the remaining feasible controls, we apply topic modeling to all their articles published between 2004 and 2021 with abstracts of at least 50 words. Using the topic model, we construct:

- $pairdist_3$: similarity between the PI and each potential match's publications in the previous 3 years. This value ranges from 0 to 1, with larger numbers representing greater similarity.
- $WAdist_3$: the weighted average of annual distances between the funded research and each scientist's publications for the three years prior to funding, as defined above, and
- $entropy_{i,3}$: breadth of topics in publications of researcher *i* in the previous 3 years.

The two distance measures consider both the similarity between the publication portfolios of the PI and proposed match, as well as the distance between each researcher's portfolio and the funded research project (measured using the funded grant abstract). Including similarity of both the PI and their match to the funded project is important, as researchers work on a range of topics and we want to ensure similarity on research topics targeted by the grant, not another area of the PIs' portfolio. To ensure similar breadth of research portfolios, the entropy of the PI and potential match in the *3* years prior to the grant must be +/- 25%.

Next, using the topic model results, we combine $pairdist_3$ and $WAdist_3$, choosing for each PI the potential match that produces the lowest $(1\text{-}pairdist_3) + WAdist_3$. To ensure that the match is a quality match, we only keep potential matches where $(1\text{-}pairdist_3)+WAdist_3 \leq 0.15$ and $WAdist_3 \leq 0.10$. These thresholds were chosen through manual inspection of a set of PIs and matched controls to determine levels of similarity at which matched researchers no longer serve as reasonable matches for the PI. Two members of the research team evaluated the top proposed matches for a selection of PIs. To ensure quality across a range of potential research topics, PIs for evaluation were chosen from each of the funding programs in our sample. After sorting pairs of PIs and proposed matches using the sum above, each pair was rated as a high quality match, low quality match, or not a match. The thresholds above were chosen to ensure that nearly all matches are high quality. We also considered rankings placing extra weight on the $WAdist_3$ term above (e.g. (1-$pairdist_3$) + 2$WAdist_3$). Finally, having chosen appropriate cutoffs, we evaluated each proposed matching rule's regression results in the years prior to the grant. Using an F-test for the joint significance of the treatment effect in the three years prior to the grant, matches determined by the lowest (1-$pairdist_3$) + $WAdist_3$ consistently yield insignificant treatment effects in the pre-treatment period. In the next section we provide a qualitative assessment of the closest and furthest match included in our sample.

We consider different matching rules to check the validity of our assumptions. Each alternate matching rule differs based on how similar the *level* and *quality* of each researcher's publication record must be. For the *level* of activity, we restrict the number of publications in the 3 years prior to the grant to be within +/- 50% or +/- 25%. In addition, some matching rules consider *quality* by also restricting the number of citations received by these publications to be within +/- 50%. In choosing between them, we trade off generating closer matches on publication quantity and quality against reductions in sample size. Our main matching rule restricts both the number of publications and citations in the 3 years prior to the grant to be within +/- 50%. We present alternate results with a larger sample by not restricting on citation counts, and on a smaller sample by limiting publication counts to be within +/- 25%.

D. Evaluation of matching

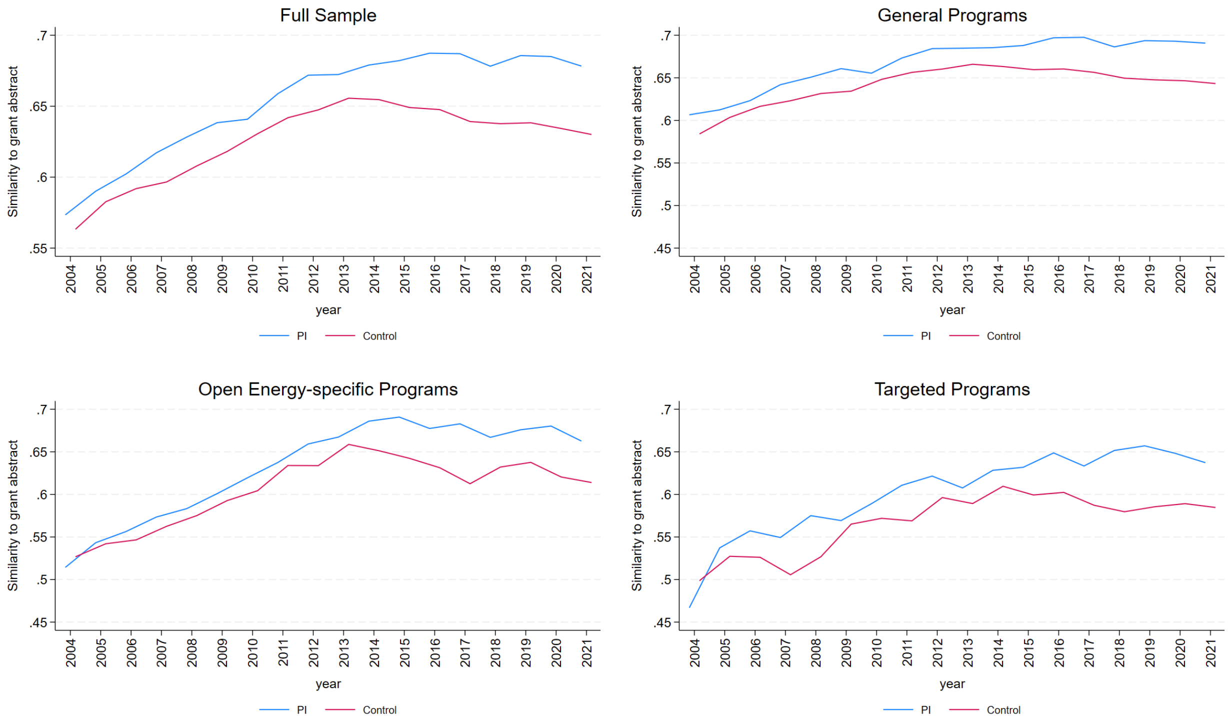
Figure D1 reproduces Figure 6 in the main text for different by start year of the grant and for different programs. Figure D2 shows how matching on topic similarity generates a set of controls who could plausibly do the work in these grants. The figure plots similarity to the grant abstract for PIs and the *entire set of feasible controls*: that is, for all researchers cited by the PIs or who cite the PIs and are similar to the PIs on other observable characteristics such as age and research institution. As we see, relative to our PIs, the research of *potential* controls is nearly half as similar to the grant topic and moves little over time, fluctuating between 0.3 and 0.35. Once we identify matched controls by also considering both (1) similarity between each PI and their potential matched controls and (2) similarity to the PI's grant, we obtain a set of controls whose research topic similarity to the grant closely tracks that of the PIs.

Table D1 compares the means of key variables for the PIs and controls in our sample. Appendix Figure D3 shows the full distribution of each variable. To check for balance in the matched sample, we present the standardized difference of means (SMD) and variance ratio (VR).[12] For an appropriately balanced sample, Rubin (2001) suggests SMD < 0.25 and VR between 0.5 and 2. Our main matching rule keeps 699 awards. The research of both PIs and controls is similarly distanced from the grant abstract (0.683 vs 0.675). Each have between 8 and 9 publications per year. While PIs publish slightly more, are cited a bit more frequently, and have slightly more co-authors, all are within the acceptable limits of an SMD less than 0.25 and VR less than 2. While limiting to publication counts within +/- 25% narrows the gap on publication counts, it does so at a cost of dropping an additional 167 awards. Similarly, not considering citations at all increases the sample size to 934 awards, but results in a VR greater than 2 for both the number of citations and co-authors (Tables D2 & D3).

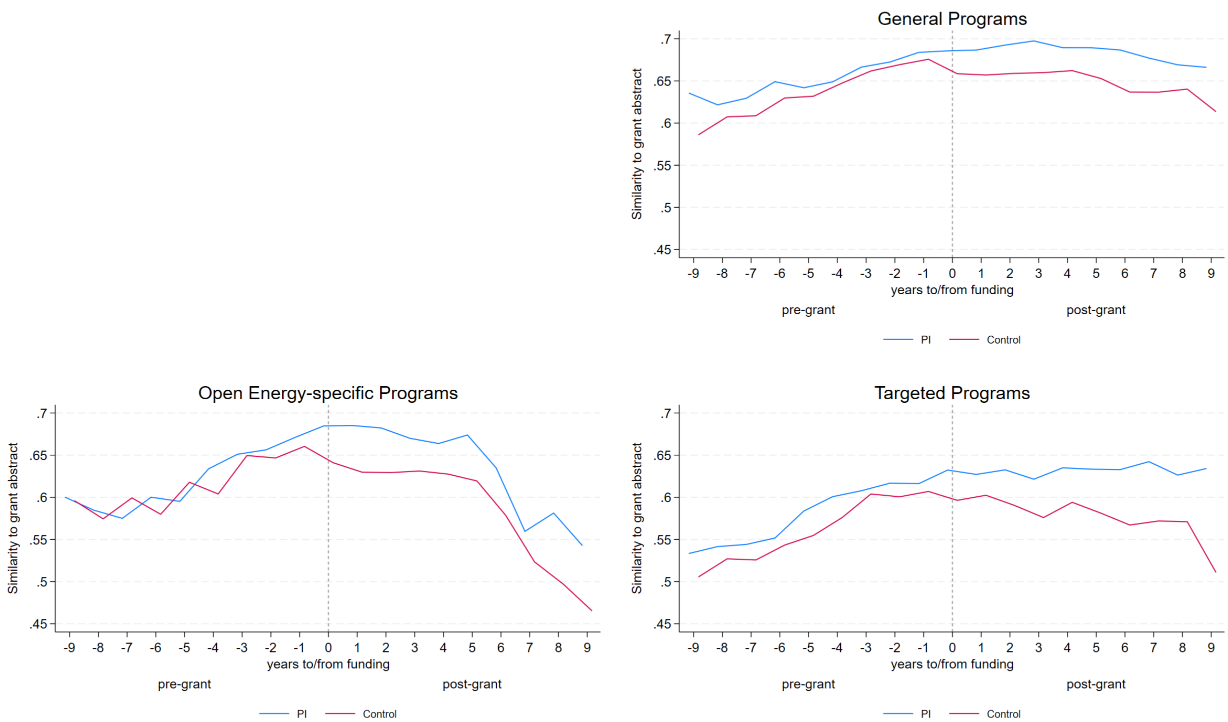
[12] Here $SMD = \frac{\bar{X}_{PI}-\bar{X}_{Control}}{\sqrt{(s_{PI}^2-s_{Control}^2)/2}}$ and $VR = {s_{PI}^2}/{s_{Control}^2}$, where $\bar{X}$ represents the sample mean and s^2 the sample variance.

Appendix Figure D1:Average Similarity to Grant Abstract

Panel A: By start year of grant

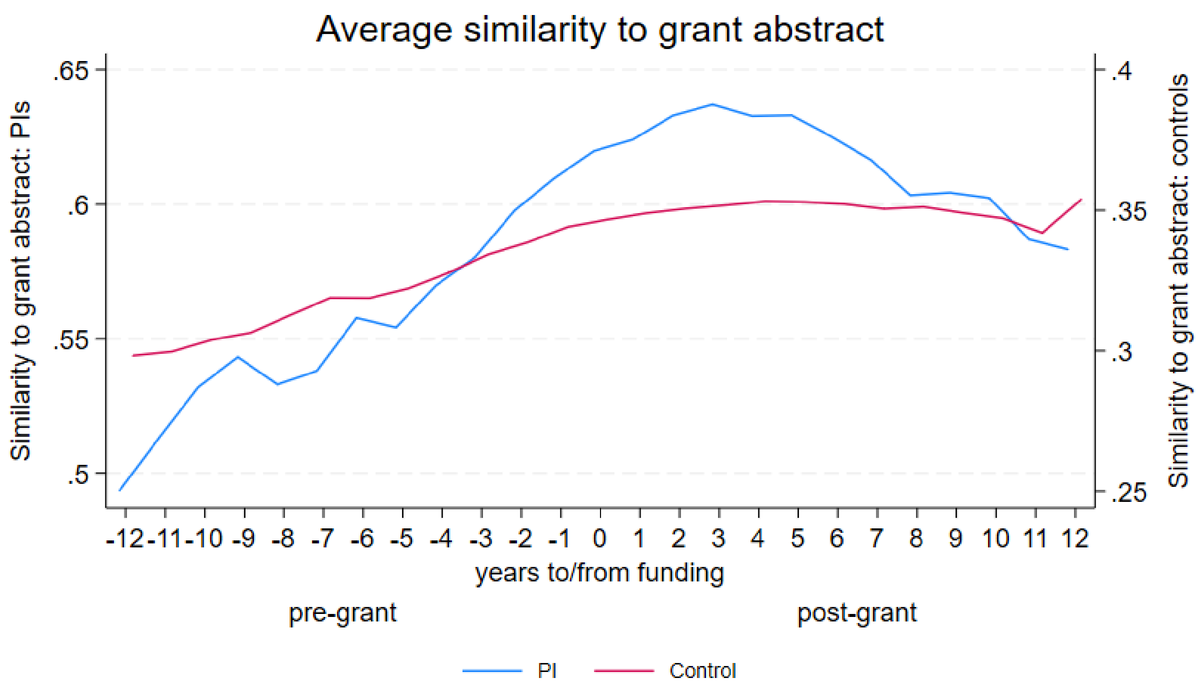


Panel B: Years to/from funding



Panel A shows the research topics of PIs (blue) and their matched controls (red) based on the start year of the grant. Panel B shows the same data relative to the start year of each grant. Both become more similar to the PI's grant abstract in the years prior to funding. Their paths diverge after funding.

Appendix Figure D2: Similarity to Grant for PIs and All Feasible Controls



NOTES: Plots the average similarity to the grant abstract for PIs and all feasible controls. To facilitate comparisons of trends across the two groups, the left and right axes have the same scale, but a different range.

Table D1: Quality of Matching, Similarity to Grant Abstract (N_{pub} +/- 50% &N_{cites} +/- 50%)

	Mean (treated)	*Mean (control)*	*SMD*	*VR*
Similarity to abstract (3 yr prior)	0.683	0.675	0.033	0.986
	(0.226)	(0.227)		
Experience at grant	22.378	20.635	0.168	1.012
	(10.430)	(10.366)		
Average citations (3 yr prior)	67.931	65.325	0.043	1.080
	(62.440)	(60.089)		
Citations (3 yr prior)	1982.674	1702.303	0.116	1.228
	(2541.985)	(2293.542)		
N co-authors (3 yr prior)	59.966	54.423	0.109	1.891
	(58.160)	(42.295)		
Entropy (3 yr prior)	1.894	1.854	0.105	1.041
	(0.382)	(0.375)		
N publications (3 yr prior)	28.322	24.913	0.184	1.263
	(19.617)	(17.453)		

# of awards: 699

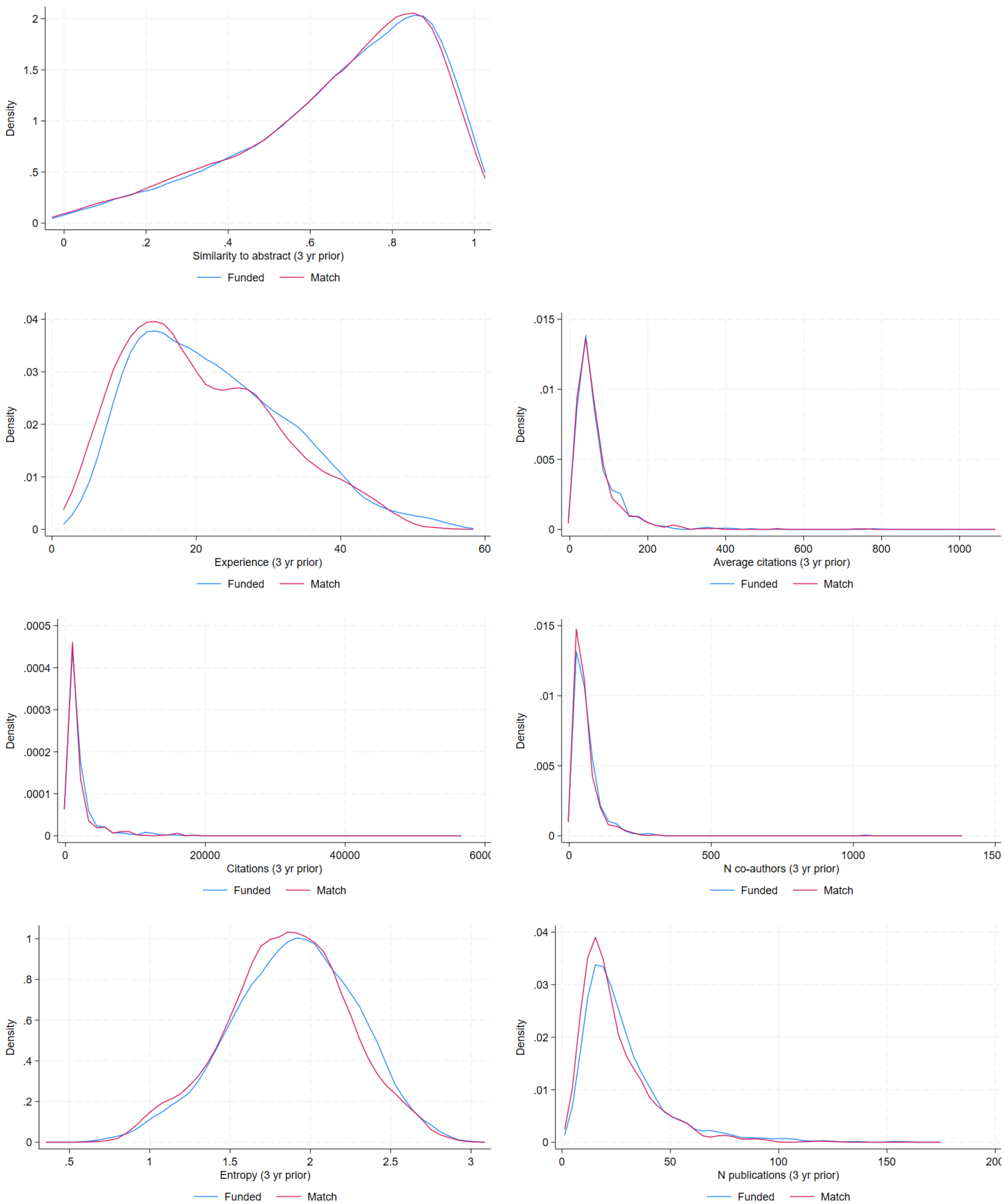
Table D2: Quality of Matching, Similarity to Grant Abstract (N_{pub} +/- 50%)

	Mean (treated)	*Mean (control)*	*SMD*	*VR*
Similarity to abstract (3 yr prior)	0.669	0.662	0.030	0.974
	(0.227)	(0.230)		
Experience at grant	21.722	19.983	0.170	1.053
	(10.358)	(10.093)		
Average citations (3 yr prior)	82.053	69.183	0.132	1.088
	(99.585)	(95.461)		
Citations (3 yr prior)	2395.451	1653.250	0.229	**2.369**
	(3849.674)	(2501.068)		
N co-authors (3 yr prior)	61.270	51.321	0.171	**3.331**
	(72.123)	(39.516)		
Entropy (3 yr prior)	1.917	1.877	0.105	1.063
	(0.391)	(0.379)		
N publications (3 yr prior)	27.929	24.030	0.215	1.413
	(19.593)	(16.481)		
Number of Awards: 934				

Table D3: Quality of Matching, Similarity to Grant Abstract (N_{pub} +/- 25% &N_{cites} +/- 50%)

	Mean (treated)	*Mean (control)*	*SMD*	*VR*
Similarity to abstract (3 yr prior)	0.696	0.686	0.041	0.991
	(0.224)	(0.225)		
Experience at grant	22.538	20.575	0.190	1.024
	(10.406)	(10.283)		
Average citations (3 yr prior)	67.971	60.170	0.139	1.463
	(61.220)	(50.609)		
Citations (3 yr prior)	2012.015	1678.233	0.132	1.462
	(2749.058)	(2273.790)		
N co-authors (3 yr prior)	57.094	56.449	0.016	1.122
	(42.650)	(40.260)		
Entropy (3 yr prior)	1.867	1.837	0.082	1.009
	(0.371)	(0.369)		
N publications (3 yr prior)	27.767	26.423	0.078	1.133
	(17.806)	(16.725)		
Number of Awards: 532				

Appendix Figure D3: Distribution of Research Characteristics, PIs and Matched Sample



Qualitative assessment of Closest and Furthest Matches

Here we provide examples of the matches generated by our algorithm. Sorting each matched pair by our ranking criterion, (1-$pairdist_3$)+$WAdist_3$, we provide a detailed description of the PI and matched researcher with the highest and lowest ranking. For each match, we present distance to the grant for both the PI ($grantdist_{PI,}$) and matched researcher ($grantdist_{Match,3}$), as well as the individual components of our matching criterion: $WAdist_3$, the weighted average of the difference between the two distances to the grant, and *pairdist3*, the distance between the research portfolios of the two researchers. We also present a random selection of ten publications from the three years prior to their grant.

<u>Closest match</u>

Award: DE-AR0000222 (2012)

Title: Transformer-Less Unified Power Flow Controller for Wind and Solar Power Transmission

Abstract: MSU is developing a power flow controller to improve the routing of electricity from renewable sources through existing power lines. The fast, innovative, and lightweight circuitry that MSU is incorporating into its controller will eliminate the need for a separate heavy and expensive transformer, as well as the construction of new transmission lines. MSU's controller is better suited to control power flows from distributed and intermittent wind and solar power systems than traditional transformer-based controllers are, so it will help to integrate more renewable energy into the grid. MSU's power flow controller can be installed anywhere in the existing grid to optimize energy transmission and help reduce transmission congestion.

$grantdist_{PI,3}$	$grantdist_{Match,3}$	$WAdist_3$	$pairdist_3$	(1-$pairdist_3$)+$WAdist_3$
0.9957	0.9941	0.00262	0.9985	0.00416

PI: Fang Zheng Peng, Michigan State University
Match: Subhashish Bhattacharya, North Carolina State University

This project, led by Fang Zheng Peng from the Department of Electrical and Computer Engineering at Michigan State University, aimed at developing a power flow controller for routing renewable electricity more affordably through power lines.

Scrutinizing their respective lists of publications confirms that our similarity metrics effectively capture qualitative similarity between the research portfolios of the PI and the match. Fang Zheng Peng's research focuses on power inverters and transmission.[13]. From 2009-2011, he published 89 papers generally related to power electronics for advanced grid integration and control (inverters, controllers, converters, transformers). Subhashish Battacharya's research within the field of power electronics is closely related, with a focus on developing technology such as Solid-State Transformers and MV power converters. His publication record for the same period comprises 77 scientific articles, some of which are related to broader issues of renewables, EV, and storage integration onto the electrical grid.

Peng representative publications (3 years prior)

- Low-Cost Semi-Z-source Inverter for Single-Phase Photovoltaic Systems
- Optimum Design Consideration and Implementation of a Parallel Hybrid Active Power Filter Integrated into a Three-Phase Capacitive Diode Rectifier
- Quasi-Z-Source Inverter with Energy Storage for Photovoltaic Power Generation Systems
- Hybrid PWM control for Z-source Matrix Converter
- Power loss analysis of current-fed quasi-Z-source inverter
- A Robust Control Scheme for Grid-Connected Voltage-Source Inverters
- A DC–DC multilevel boost converter
- Z-Source-Converter-Based Energy-Recycling Zero-Voltage Electronic Loads
- A Family of Z-source and Quasi-Z-source DC-DC Converters
- 3X DC-DC Multiplier/Divider for HEV Systems

Bhattacharya representative publications (3 years prior)

- Smart Vehicles in the Smart Grid: Challenges, Trends, and Application to the Design of Charging Stations
- Series Connected IGCT Based Three-Level Neutral Point Clamped Voltage Source Inverter Pole for High Power Converters
- A New Nine-Level Active NPC (ANPC) Converter for Grid Connection of Large Wind Turbines for Distributed Generation
- Investigation of High Performance Heat Sink Characteristics in Forced Convection Cooling of Power Electronic Systems
- A simplified space vector based current controller for any general N-level converter
- Modeling of the impact of diode junction capacitance on high voltage high frequency rectifiers based on 10kV SiC JBS diodes
- A Discrete Matlab–Simulink Flickermeter Model for Power Quality Studies
- Active Power Transfer Capability of Shunt Family of FACTS Devices Based on Angle Control
- Testing of a controller for an ETO-based STATCOM through controller hardware-in-the-loop simulation
- Control Strategies for Battery Energy Storage for Wind Farm Dispatching

Furthest match included in sample

Award: DE-SC0016367 (2016)

Title: Single-Site Metal Organic Complexes on Oxide Supports for Selective Alkane Functionalization

[13] https://grid.pitt.edu/people/fang-zheng-peng, accessed August 5, 2025.

Abstract: Selective conversion of small alkane molecules into economically-useful chemical feedstocks and fuels is a critical problem in the field of chemical catalysis. This problem is made more acute by the relative abundance of natural gas, rich in methane, and the increasing need to develop more environmentally friendly uses for our limited hydrocarbon resources. Selective functionalization will enable chemistries such as direct methane oxidation to methanol and oxidative coupling of methane to form molecules that can be used industrially for materials synthesis and chemical production. Each of these relies on the ability to selectively activate specific chemical bonds in the alkanes. This research project is a study of a metal-organic complexation method to develop chemically uniform transition metal catalyst centers on oxide surfaces. Single-site centers are known to effect selective alkane chemistries in natural systems and in a few industrial processes. It is broadly recognized that combining the catalytic properties of metal-organic complexes with those of high dispersion oxide support materials is essential to advancing next-generation catalysis. This objective falls within the mission of the Catalysis Science program and is addressed in this program through parallel studies of catalyst sites in highly-controlled model oxide systems and of high surface area catalysts in operating conditions.

$grantdist_{PI,3}$	$grantdist_{Match,3}$	$WAdist_3$	$pairdist_3$	$(1\text{-}pairdist_3)+WAdist_3$
0.8164	0.7967	0.04797	0.89883	0.14914

PI: Steven Tait, Indiana University
Match: Thomas Senftle, Penn State

This project, let by Steven Tait of Indiana University, studies chemical feedstocks with the goal of developing environmentally friendly chemical applications. While this pair is the lowest ranked of any pair in our sample, their work in the three years prior to funding appears suitably similar. While not obvious from the titles, the abstracts of both scientists' work describe research on chemical catalysis.

Tait representative publications (3 years prior)

- Multifunctional Tricarbazolo Triazolophane Macrocycles: One-Pot Preparation, Anion Binding, and Hierarchical Self-Organization of Multilayers
- Two- and Three-Electron Oxidation of Single-Site Vanadium Centers at Surfaces by Ligand Design
- Self-assembling Sierpiński triangles
- Living on the edge: Tuning supramolecular interactions to design two-dimensional organic crystals near the boundary of two stable structural phases
- Redox-active on-surface polymerization of single-site divalent cations from pure metals by a ketone-functionalized phenanthroline
- Selective Anion-Induced Crystal Switching and Binding in Surface Monolayers Modulated by Electric Fields from Scanning Probes
- High-Fidelity Self-Assembly of Crystalline and Parallel-Oriented Organic Thin Films by π–π Stacking from a Metal Surface
- Anion-induced dimerization of 5-fold symmetric cyanostars in 3D crystalline solids and 2D self-assembled crystals
- Redox-Active On-Surface Assembly of Metal–Organic Chains with Single-Site Pt(II)
- Interfacial Organic Layers for Chemical Stability and Crystalline Ordering of Thiophene and Carboxyl Films on a Metal Surface

Senftle representative publications (3 years prior)

- Charge Transfer Stabilization of Late Transition Metal Oxide Nanoparticles on a Layered Niobate Support
- Role of Site Stability in Methane Activation on Pd x Ce1–x Oδ Surfaces
- Molecular Dynamics Investigation of the Effects of Tip–Substrate Interactions during Nanoindentation
- Illuminating surface atoms in nanoclusters by differential X-ray absorption spectroscopy
- Determining in situ phases of a nanoparticle catalyst via grand canonical Monte Carlo simulations with the ReaxFF potential
- Influence of Hydroxyls on Pd Atom Mobility and Clustering on Rutile $TiO_2(011)$-2 × 1
- A ReaxFF Investigation of Hydride Formation in Palladium Nanoclusters via Monte Carlo and Molecular Dynamics Simulations
- Development of a ReaxFF potential for Pd/O and application to palladium oxide formation

E. Results Removing Papers Acknowledging Grants

One possible concern is that our finding that funded researchers moving in the direction of their grant is purely mechanical – that their subsequent publications are similar to their funded grant abstract by construction. Here we drop publications that acknowledge the funded award. We match award numbers to paper acknowledgements in the Dimensions database for the PIs in our sample. We search the acknowledgements of each PI's papers for award numbers matching the awards in our data (after removing punctuation such as hyphens) and use additional acknowledgement text to verify that each matched award numbers belongs to an award made by the DOE or NSF. Since we are unable to match all awards to at least one acknowledging paper, we drop awards without any acknowledgements from our sample.[14] Figure E1 shows three sets of results using the differences-in-differences specification. Blue lines are the original results from the main text. Red lines repeat our main analysis on a restricted sample that drops awards never receiving acknowledgements. Because these results are similar, any differences observed when dropping papers acknowledging grants from the sample is not simply due to changing sample sizes. Green lines show the results of regressions excluding papers that directly acknowledge the PI's grant. Figure E2 presents the event study results, omitting differences between the main specification with and without awards that never receive acknowledgements, since as in Figure E1 they are nearly identical.

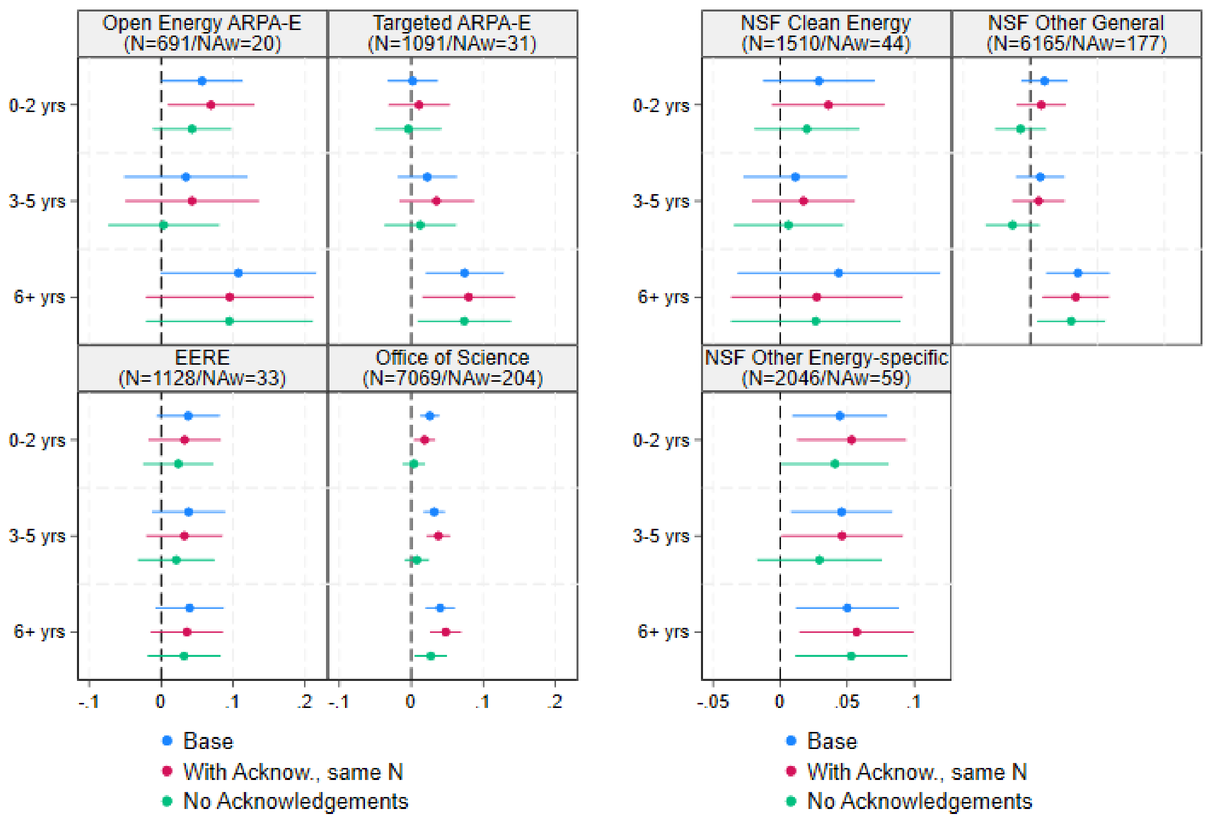
As would be expected, we observe less movement towards the grant topic when dropping papers directly acknowledging the grant. For general award programs, there is no short-run movement towards the funded topic after removing papers acknowledging the grant. This suggests that these researchers use general awards to continue research on topics for which they have previous experience. For these researchers, the effect of funding on research direction goes to zero when removing papers acknowledging the grant (Figure E2) before returning to a long-run upward trend. In combination, general funding programs appear to help PIs continue established research directions

For both EERE and NSF energy-specific awards, publications of PIs are more similar to the grant

[14] To verify that our matching doesn't underidentify acknowledging papers, we also did robustness checks dropping awards not receiving at least 3 or at least 5 acknowledgements. Results are similar to those presented here.

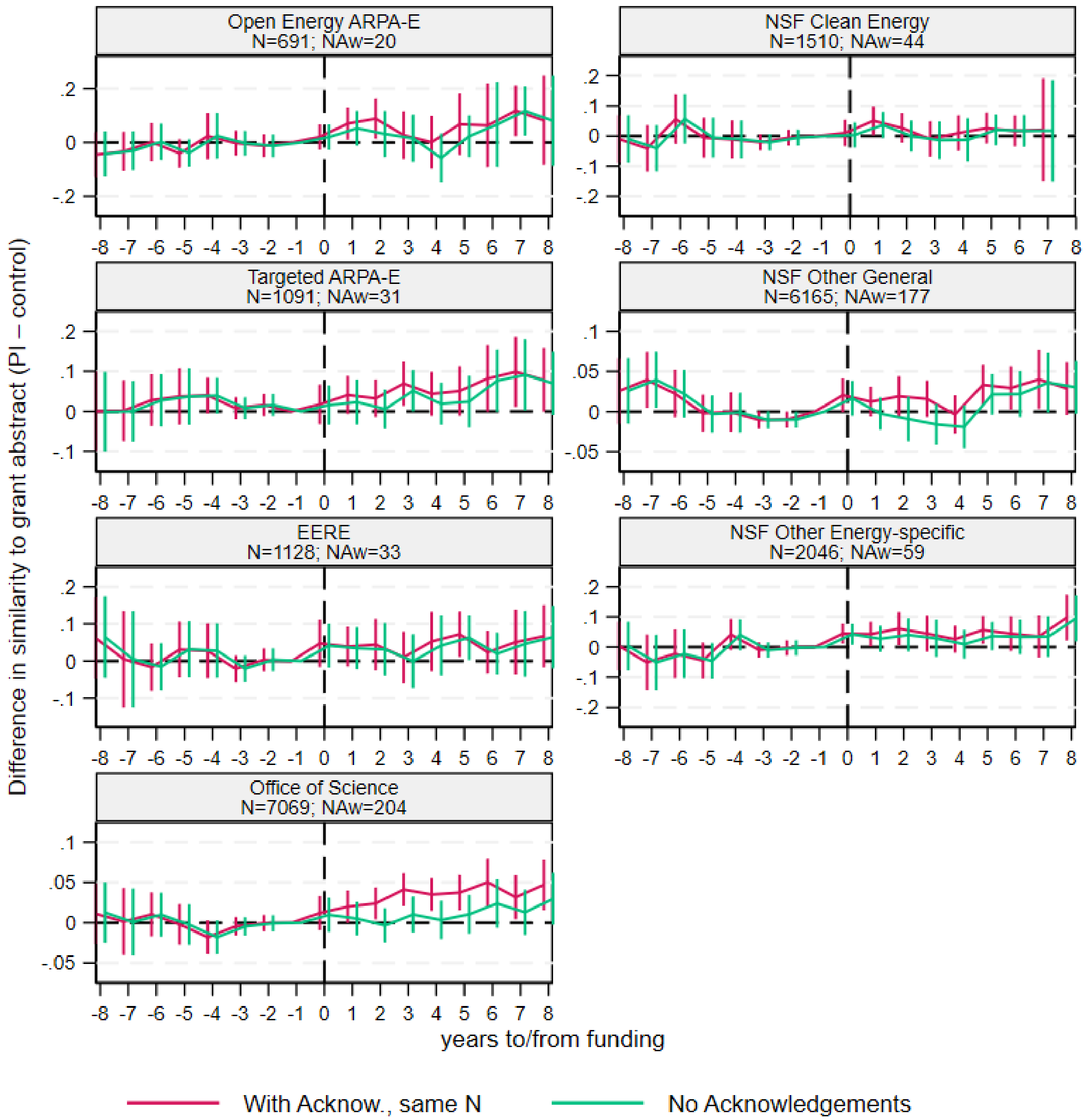
abstract even *in the year of funding* (Figure E2, EERE: β=0.0484, p-value 0.095; NSF β=0.047, p-value 0.017). Along with finding less change when removing acknowledgement for these programs, these programs appear to fund work consistent with the overall research portfolio of PIs. In contrast, for ARPA-E we observe less movement towards the grant topic when removing papers acknowledging the grant, but not no movement. Moreover, the persistent long-run effects of ARPA-E suggest that bringing scientists into the energy space develops human capital. In sum, general awards seem to mainly help research already underway, but ARPA-E funding tends to shift researchers into new territory, which shows up even in papers that never acknowledge the grant. Moreover, the grants appear to create durable expertise rather than being one project output, which is consistent with the agency's mission.

Figure E1: Difference-in -difference results removing papers acknowledging grant



NOTES: N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

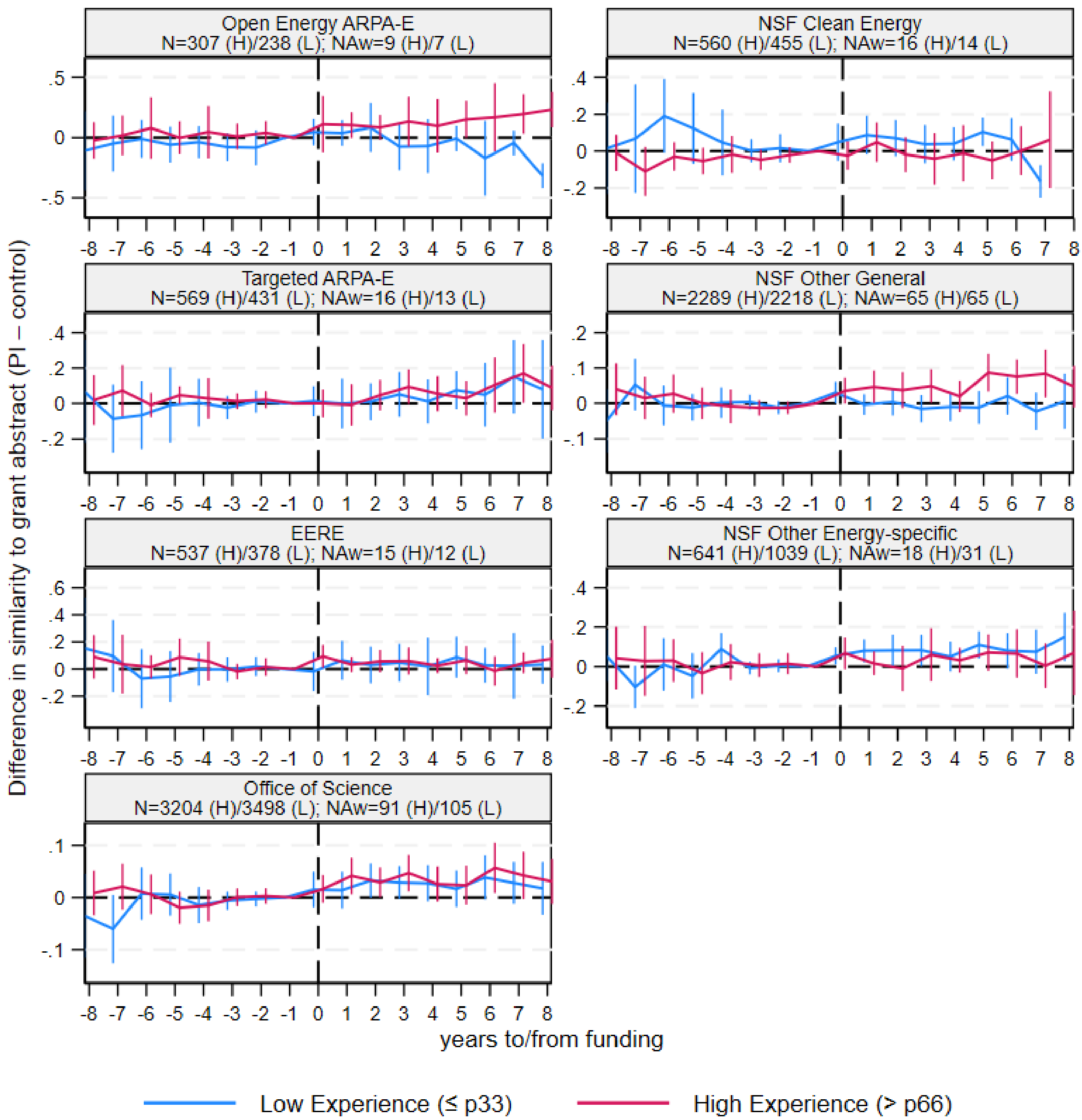
Figure E2: Event study results removing papers acknowledging grant



NOTES: N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

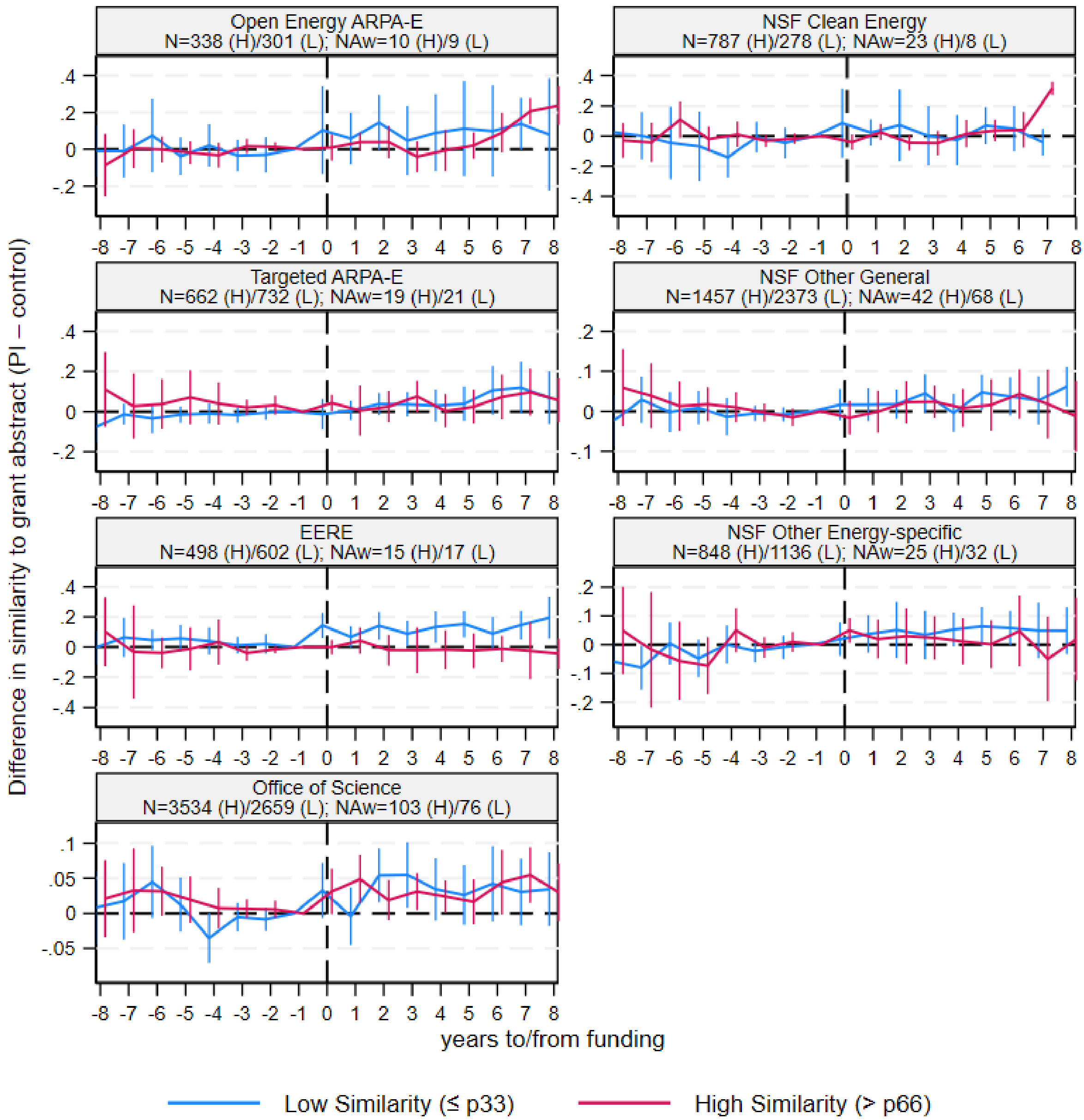
F. Results by PI Characteristic

Figure F1: Effect of experience on similarity



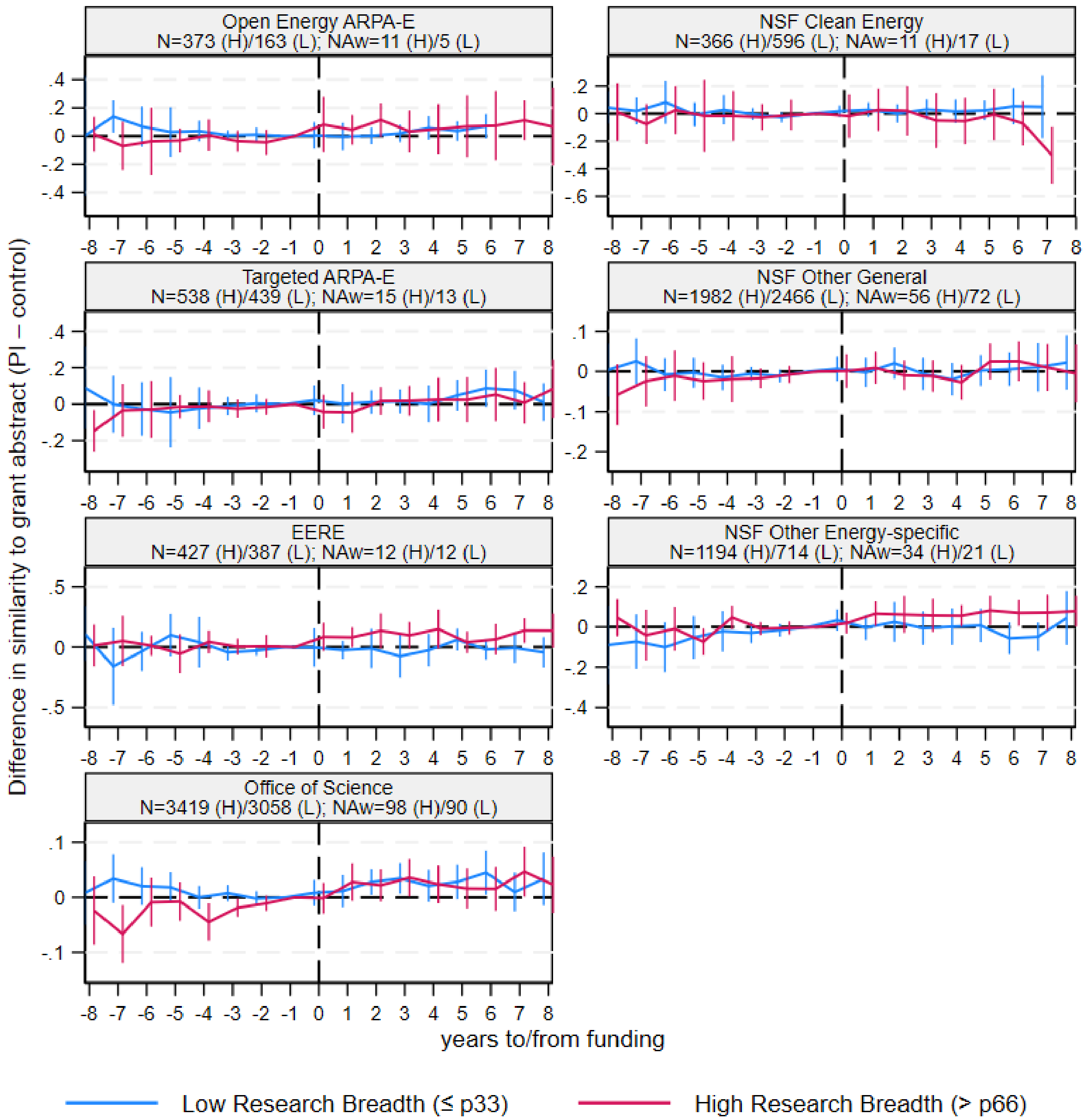
NOTES: Results of event study regressions comparing the similarity of a PI's publications to their grant abstract, relative to their matched control. N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

Figure F2: Effect of prior similarity to grant



NOTES: Results of event study regressions comparing the similarity of a PI's publications to their grant abstract, relative to their matched control. N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

Figure F3: Effect of research breadth on similarity



NOTES: Results of event study regressions comparing the similarity of a PI's publications to their grant abstract, relative to their matched control. N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

G. Robustness Checks

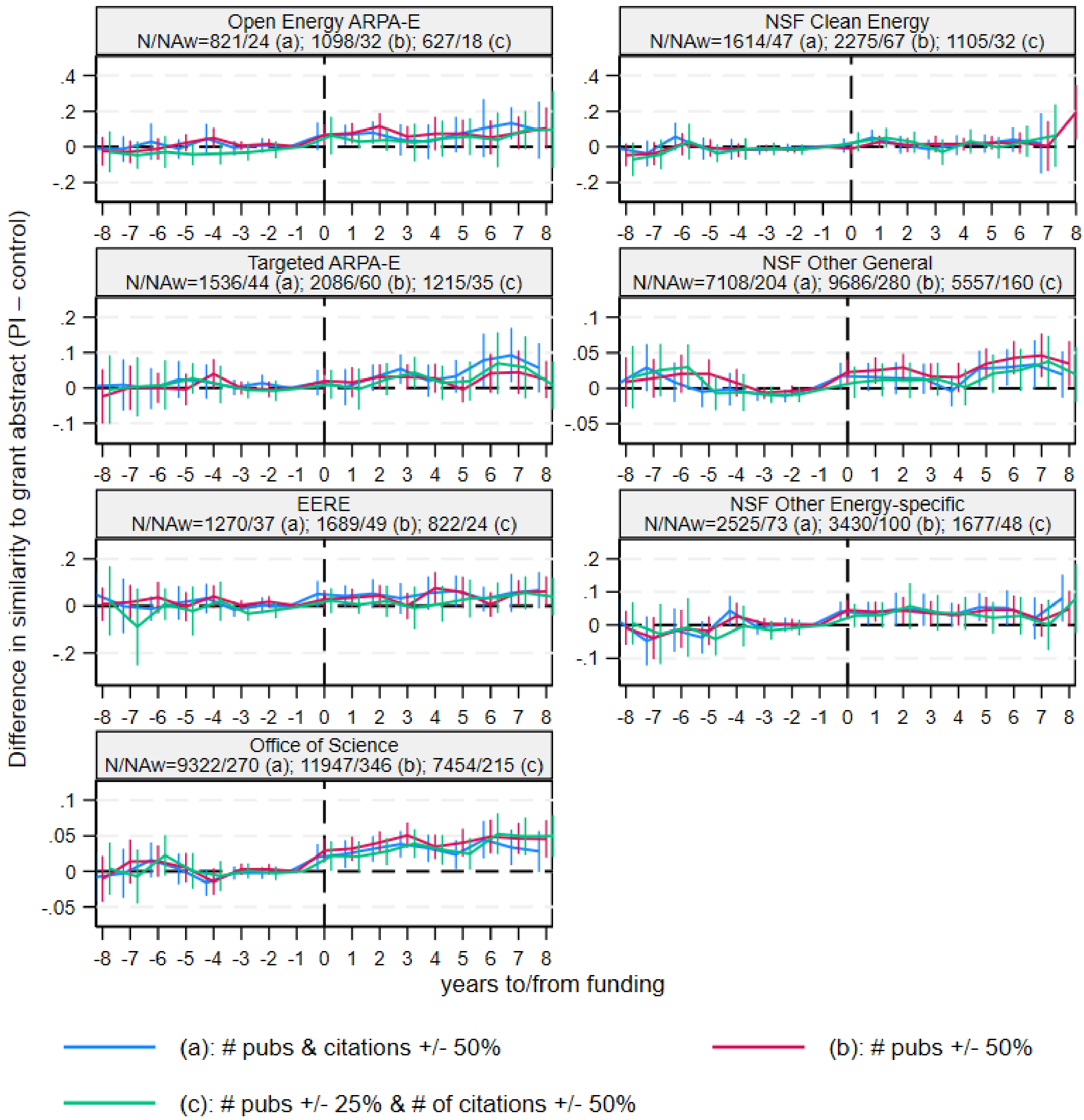
In this appendix we provide additional details on the robustness checks briefly described in the results section of the main paper.

i. Robustness to different matching rules

In the main text, we require that both the number of publications and citations in the 3 years prior to the grant for both PIs and potential matches be within +/- 50%. Here we consider robustness to two alternative cutoffs. Not considering citations at all increases the sample size to 934 awards, but as noted in the text, results in a VR greater than 2 for both the number of citations and co-authors. Limiting publication counts to a range within +/- 25% provides a closer match on publication counts, but at a cost of dropping an additional 167 awards.

Figure G1 compares the results using these alternative rules (in red and green) to our main specification (in blue). In general, our results are robust. Only considering publication counts results in a slightly higher treatment effect for the full sample, but also in a larger difference between the funded PIs and their matches in the pre-treatment period. The smaller sample size of the more restrictive matching rule provides similar results for the full sample, but smaller effects for some of the smaller programs in our data, such as ARPA-E.

Figure G1: Robustness to Different Matching Rules, All Programs



NOTES: N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

ii. Removing limitations on closeness of PIs and controls

Limiting our matched sample to control scientists whose research is sufficiently close to the PI ensures internal validity. However, this limits external validity if the matched sample differs from the full sample of PIs. Compared to the full sample, the PIs who remain in the sample after matching tend to be more similar to the funded research (Figure G2). This limitation is most notable for programs such as ARPA-E that attract researchers working on a diverse range of topics. In particular, for ARPA-E, we are most likely to find a match for both PIs whose prior work is similar to the grant and those whose work is very different from the grant. However, when restricting our matches to those "sufficiently close" to one another, we are unable to find matches for some PIs, especially in the middle of the grant similarity distribution. While the prior research of most ARPA-E PIs is only moderately similar to their grant abstract (e.g., similarity between 0.4 and 0.6), that is not the case for the matched sample (Figure G2). Intuitively, finding a match for someone who spends half of their time working on energy and half on another topic is more difficult than finding matches for researchers who spend nearly all their time or very little time working on energy prior to funding.

Similarly, while other pre-grant characteristics of the PIs in our matched sample are generally similar to the full sample, there are a few exceptions. For example, NSF Clean Energy researchers in the matched sample are older, and Targeted ARPA-E PIs in the matched sample have lower entropy than the full sample (Figure G3). Therefore, ensuring external validity requires examining our proximity filter more closely.

To assess the external validity of our results, we re-run our analysis on a different matched sample that doesn't require the control researcher to be "sufficiently close" to the PI. We maintain all other matching requirements, but simply take the control with the most similar research profile, whether or not it is "sufficiently close". We do this in one of two ways: (1) *Closest weighted sum*: choosing the match with the lowest (1-$pairdist_3$) + $WAdist_3$, without the restrictions that (1-$pairdist_3$)+$WAdist_3 \leq 0.15$ and $WAdist_3 \leq 0.10$; (2) *Closest pair similarity*: ignoring similarity to the grant proposal, and simply choosing the control with the smallest $pairdist_3$, again without restricting the pair to be sufficiently close.

Figure G2: Distribution of similarity to the funded research, 3 years prior to grant

Full sample *Matched Sample*

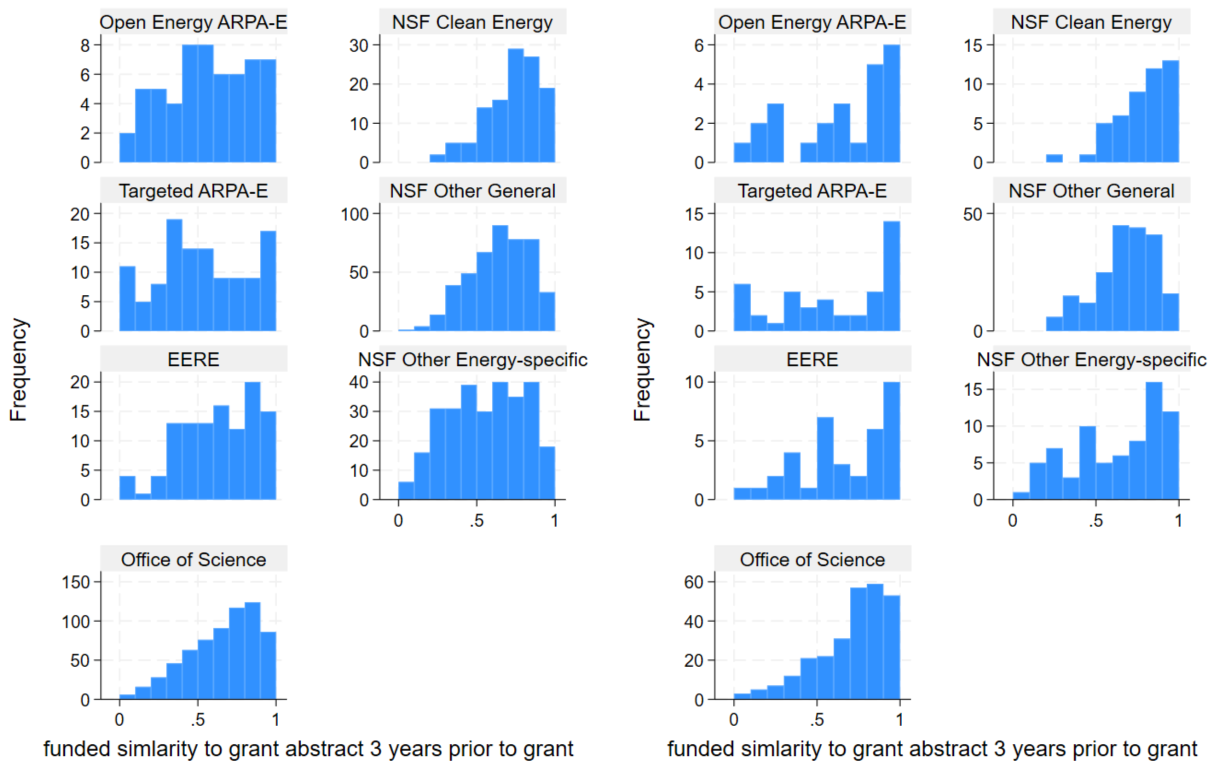
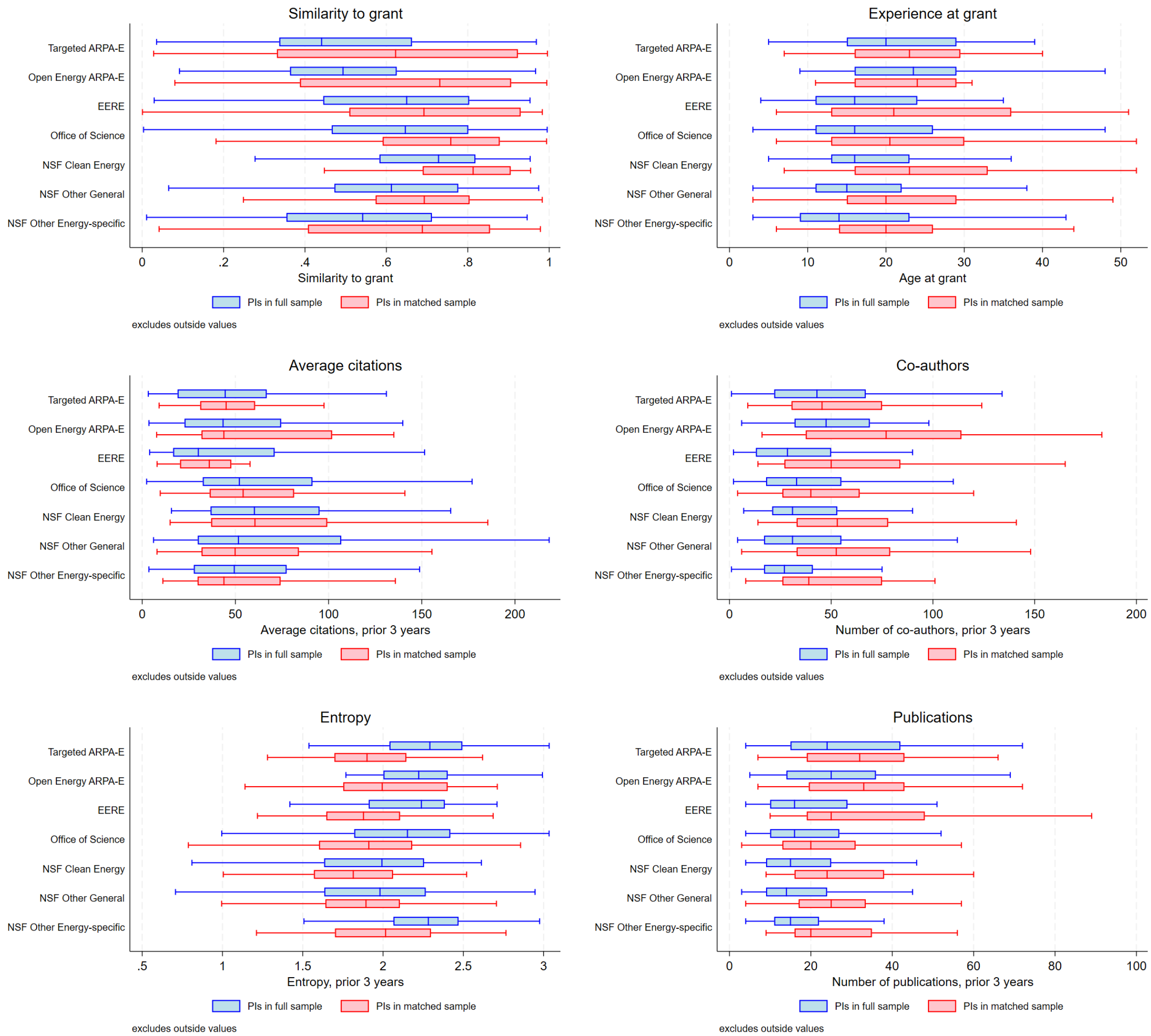
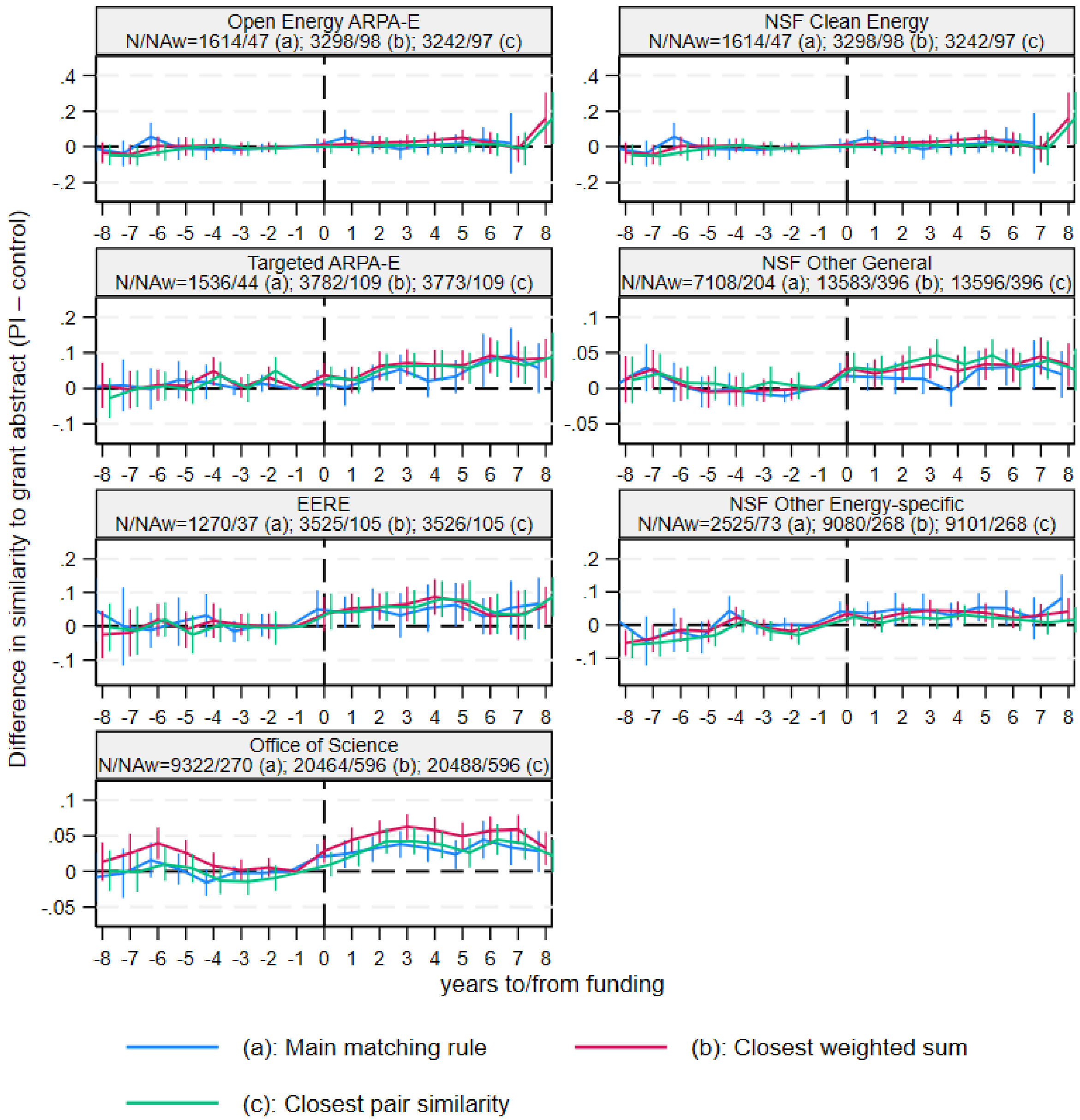


Figure G3: Distribution of PI characteristics, 3 years prior to grant



NOTES: Boxplots excluding outliers for clearer visualization of central tendencies. Light blue boxes represent PIs in full sample; light red boxes represent PIs in matched sample across different program types. The box represents the interquartile range, with the middle line indicating the median value. The whiskers show the range of the data using upper and lower adjacent values, omitting outliers.

Figure G4: Robustness to Removing Matching Filters



NOTES: N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

Figure G4 compares results using these alternative samples to our main results. Blue lines present our original results, and red and green lines the alternative matching rules. In general, the patterns are similar, but with larger post-funding changes using the less restrictive samples. However, we also observe cases where the research of the PIs and control differ prior to funding (such as for DOE's Office of Science). Notably, while differences in the distribution of similarity before and after matching were most notable for ARPA-E, the results for ARPA-E appear robust. Even with the alternative matching rules, we observe no significant differences between PIs and controls before funding, and observe large differences after funding.

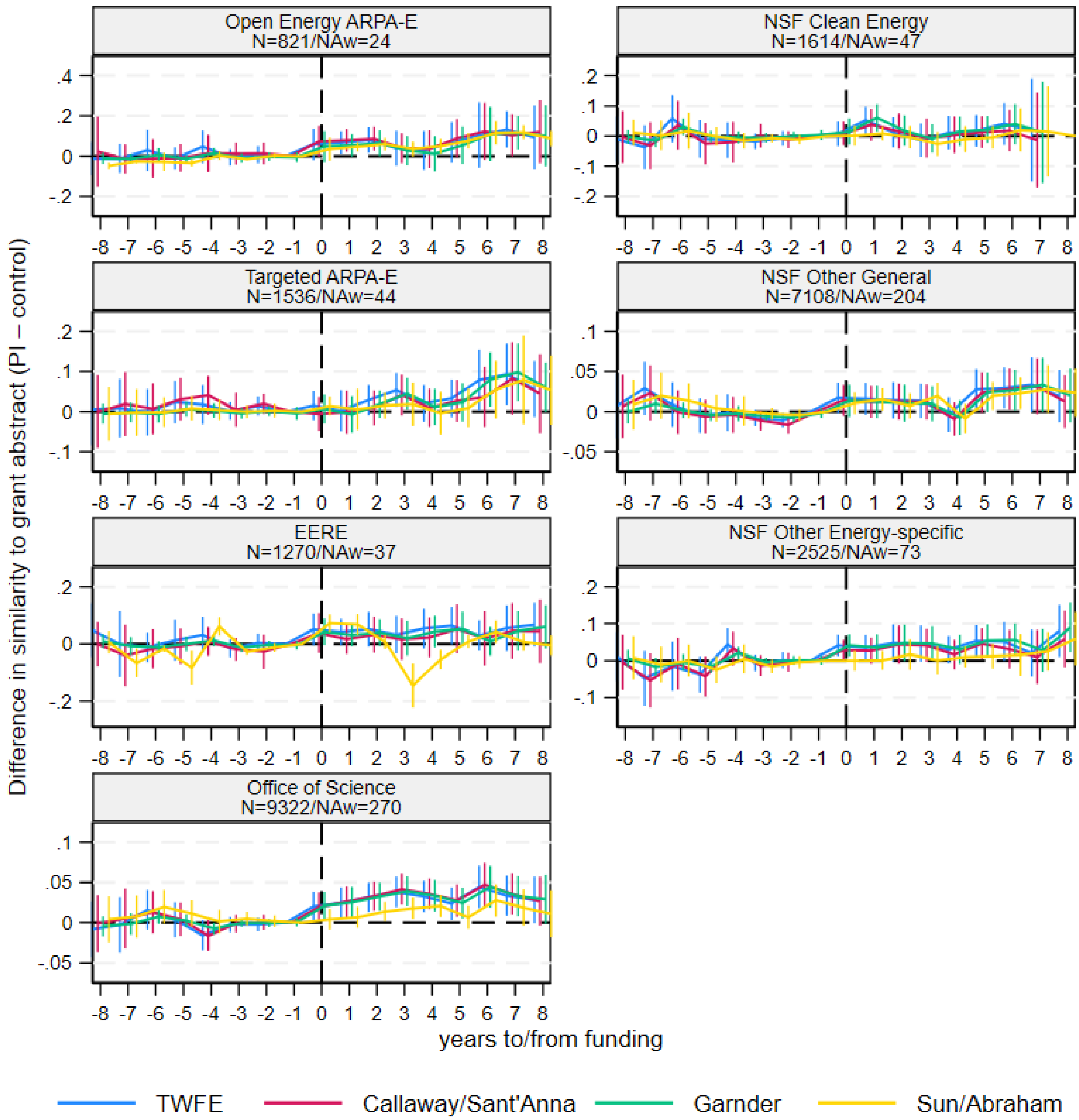
iii. Alternative Estimation Techniques

Our main models use both program and cohort-specific year fixed effects. As described in the main text, cohort-specific year fixed effects ensure the regressions only make valid comparisons across pairs of researchers. In this section we provide further documentation of the importance of cohort-specific year fixed effects, as well as demonstrating robustness to new two-way fixed effect (TWFE) estimation techniques. The important takeaway is that it is the application of different year fixed effects, rather than different estimation techniques, that affect the results.

This is illustrated by Figures G5 and G6. Both figures present results using different TWFE estimation techniques. Blue lines come from a standard TWFE estimation. Red lines apply the TWFE estimator in Callaway and Sant'Anna (2021). Green lines apply the two-stage estimator from Gardner (2022). Yellow lines present results using the estimator proposed in Sun and Abraham (2020).

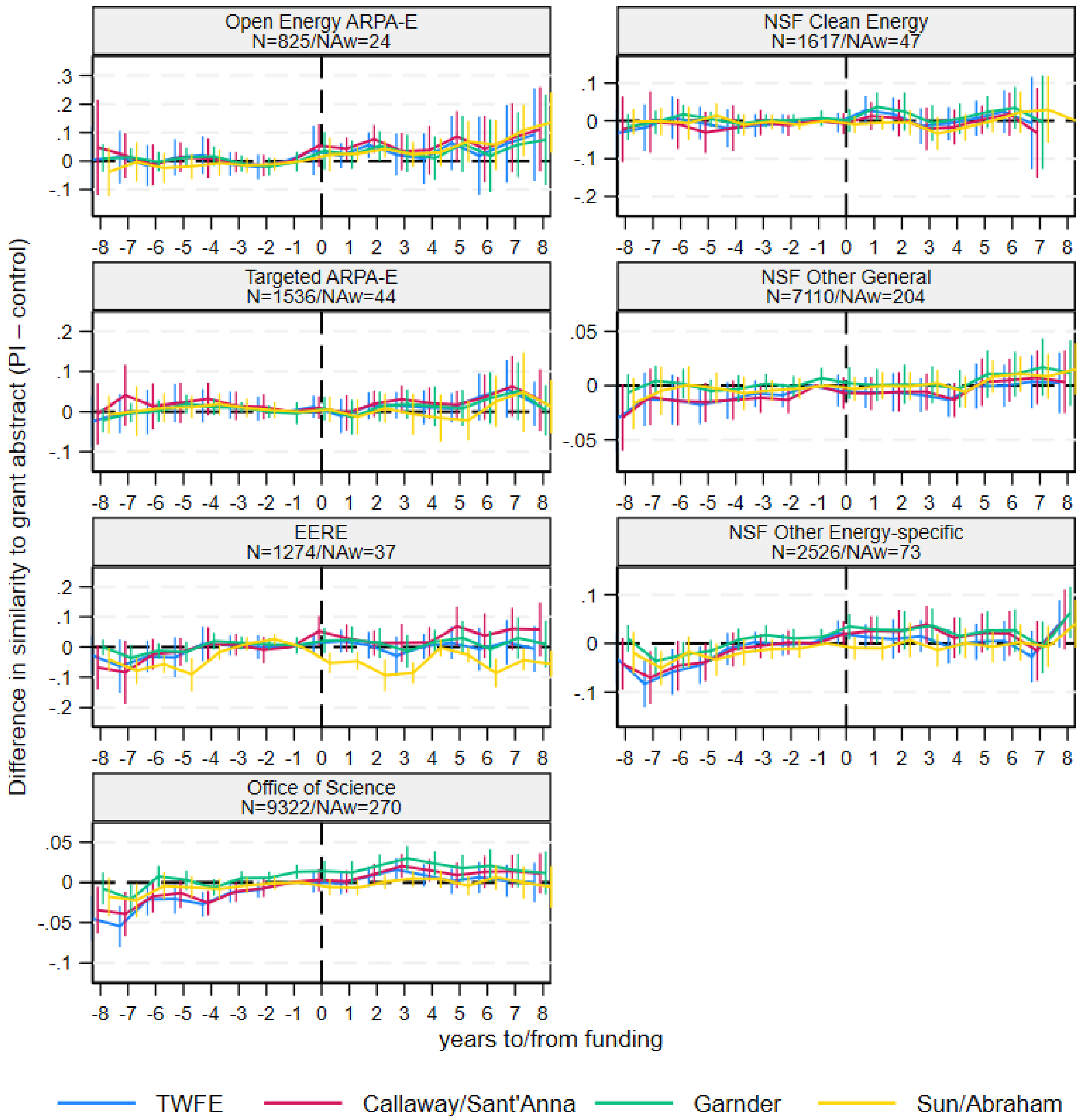
Figure G5 uses our preferred specification including program/cohort-specific year fixed effects. In general, our results are robust to different estimation techniques. The one exception is slightly smaller coefficients using Sun and Abraham's estimation, particularly for general programs.

Figure G5: Robustness for Estimation Methods – Year X Program X Cohort Fixed Effects



NOTES: N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

Figure G6: Robustness for Estimation Methods – Year X Program Fixed Effects



NOTES: N = # of observations. NAw = # of awards. All models include individual/award fixed effects and controls. Standard errors clustered by award. Note different scale of y-axes.

Similarly, when only using program-specific year fixed effects (Figure G6), the TWFE estimation technique chosen makes little difference. While the results using Gardner's method are shifted upward slightly, that occurs here because they do not require normalizing the coefficient in year *t*-1 to equal 0. However, compared to our baseline model with program/cohort-specific year fixed effects, some results are different. In some cases, PIs are moving in the direction of their funded research even before funding: their research begin less similar to the grant than the controls but approach the grant topic as funding nears. In contrast, we see no evidence of pre-trends for either ARPA-E program. Thus, the result that ARPA-E both attracts new researchers to clean energy and moves their research in new directions is robust to the choice of year fixed effects.[15]

Importantly, the negative pre-funding estimates are most prominent for the Office of Science, NSF Other General, NSF Other Energy-Specific programs. These programs give PIs more flexibility, and thus will attract a broader range of topics. Thus, the set of awards in a given year may be very different from the set of awards in a different year. These differences may be due to changes in scientists' interests over time, but may also occur because of funding allocation changes. For example, in select years, NSF made additional funds available for clean energy across multiple directorates. Proposals were to be submitted to the appropriate NSF disciplinary area. We code awards made in response to these calls, which are identified in NSF's funding database with reference code 8396, as Open Energy-specific awards. Such awards don't appear in our database until 2012, and most occur in 2015 and 2016. Thus, clean energy funding proposals made to the "general" NSF programs Chemical Catalysis and Electronic & Photonic Materials Programs would be considered a general award prior to 2012. However, awards to the same programs in response to the additional clean energy funding available after 2012 would be considered Open energy-specific awards if they contain reference code 8396. The remaining General awards in these years would be less likely to include projects related to clean energy. As a result, assuming a common year fixed effect for these programs across different years would be incorrect.

[15] Note that all the results presented in Figure F6 are nearly identical if only using year fixed effects.

iv. Robustness of citation analysis

In the main manuscript, we use the number of citations received by scientific articles within a five-year window as the dependent variable in our analysis of the relationship between similarity to the grant and research impact. Because papers from older cohorts have more time to accrue citations, this five-year window ensures that we sum up citations in a consistent way across cohorts. Five years is generally sufficient to allow time for citations to peak. However, this measure does not take into account differences in citation trajectories across fields of science. To mitigate this, we estimate effects separately for different NSF and DOE funding programs. However, this can only partially control for differences across fields, as even within each program we observe heterogeneity in the fields of the papers acknowledging the grants. This could bias our estimates if PIs in different regions of the similarity distribution (those whose prior research is highly dissimilar, or highly similar, to the grant), pivot away from fields that are markedly more, or less, cited than clean energy research. To provide reassurance that our results are not affected by this concern, here we show robustness to using two alternative measures of the citation variable that adjust for differences across fields.

In Figure G7, we show the marginal effects of similarity on the papers' Field Citation Ratio (FCR)[16]. The FCR normalizes citation rates by dividing the number of citations ever received by the average number of citations received by papers published in the same field and year. We extracted FCR values from Dimensions, but those are missing when an article is only associated with 2-digit FoR codes or when the 4-digit FoR category contains fewer than 500 papers in the publication year.[17] These missing values explain why the number of observations reported in Figure G8 differs from Figures 9 & 10.

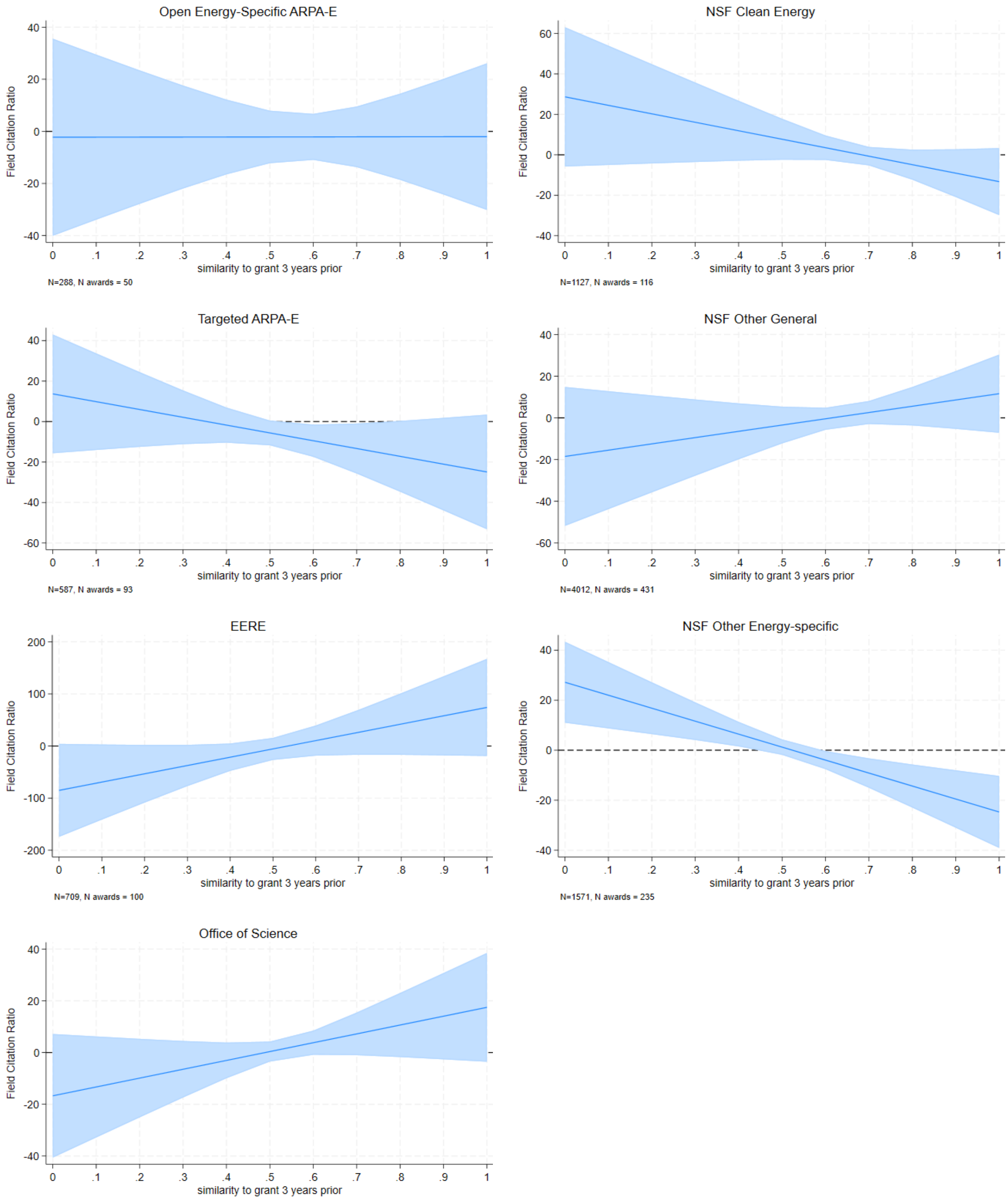
Second, following Hill et al (2025), we identify publications that are in the upper 5% of the citation distribution for their field. Results from this model are presented in Figure G8. Similar to the FCR, this dummy variable adjusts for differences in citation rates across fields and publication years. However, the interpretation here changes slightly: Figure G8 specifically informs on whether PIs who are (dis)similar to

[16] Note that we also use the FCR measure to control for each PIs' average citation rate in the pre-period.
[17] https://dimensions.freshdesk.com/support/solutions/articles/23000018848-what-is-the-fcr-how-is-it-calculated-

their grant are more likely to produce research that is at the very top of the field in terms of citation impact. Again, our results are robust when using this alternative dependent variable.

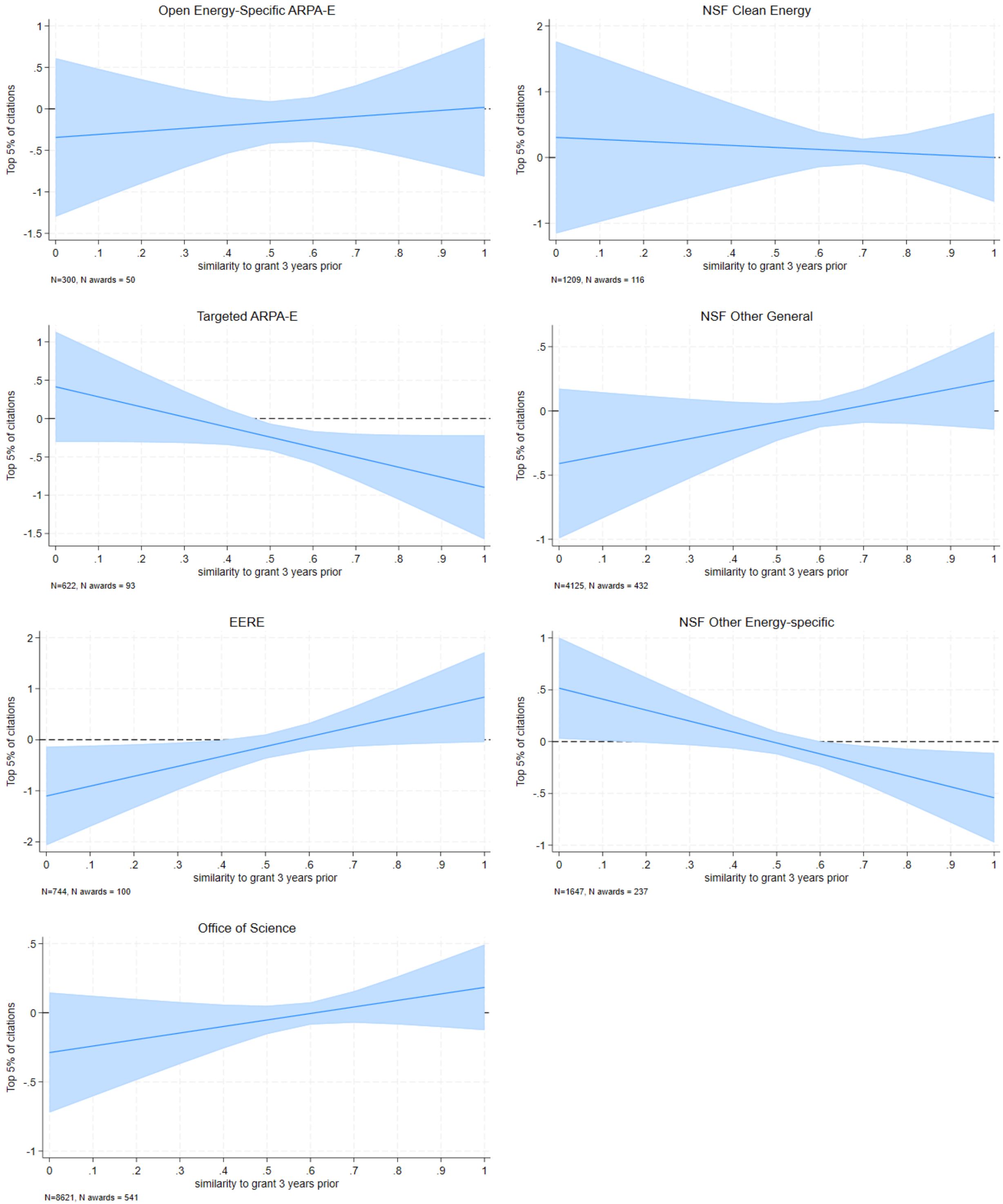
Finally, Table G1 provides additional support for our use of a quadratic model. Here we compare the results of our quadratic model to a model using only a linear term for similarity to the grant abstract. Whether the squared term adds new information is program specific. The squared term is only statistically significant for the Targeted ARPA-E, EERE, and NSF Other Energy-specific programs. Consistent with Figure 10, these are the programs where the quadratic relationship is most notable. As collinearity between the linear and squared term may reduce statistical significance of each coefficient, we also present the p-value from a joint test of significance of both the linear and squared similarity terms. Only for Targeted ARPA-E (p-value 0.052) and NSF Other Energy-specific (p-value 0.004) are these coefficients jointly significant. However, for the remaining programs, the interpretation of results is the same using either a linear or quadratic model. In neither case is the direct effect of similarity to the grant even close to statistically significant. Moreover, for the Targeted ARPA-E and NSF Other Energy-specific programs, the coefficients on similarity to the grant are negative but imprecisely estimated using the linear model. Thus, the quadratic model correctly fits the data for these programs, and does not incorrectly obscure any linear effects that might be observed in other programs.

Figure G7: Marginal effects of similarity on citations using Field Citation Ratio (FCR)



NOTES: Graphs show marginal effects. All models control for # of publications & citations in prior 3 years, age at grant, and publication year fixed effects. Standard errors are clustered by award number. Shaded region shows 95% confidence intervals.

Figure G8: Marginal effects of similarity on citations, using top 5% of citations



NOTES: Graphs show marginal effects. All models control for # of publications & citations in prior 3 years, age at grant, and publication year fixed effects. Standard errors are clustered by award number. Shaded region shows 95% confidence intervals.

Table G1: Effect of similarity on citations, linear vs. quadratic model

	Targeted ARPA-E		Open ARPA-E		EERE		Basic Energy Sciences	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Similarity to grant	-23.164	121.111	-41.694	-19.745	47.686	-405.649*	1.787	-82.097
	(16.15)	(78.77)	(28.25)	(101.65)	(72.27)	(222.19)	(10.14)	(60.91)
(Similarity to grant)^2		-141.844*		-20.387		377.814*		71.329
		(75.71)		(84.06)		(213.01)		(57.05)
N	648	648	309	309	781	781	9396	9396
R-squared	0.162	0.171	0.15	0.15	0.053	0.078	0.107	0.108
p-value: joint sig. of sim.	0.155	0.052*	0.146	0.274	0.511	0.194	0.86	0.311

	NSF Clean Energy		NSF Other General		NSF Other Energy-specific	
	(1)	(2)	(3)	(4)	(5)	(6)
Similarity to grant	0.308	62.894	5.275	-92.299	-5.379	131.209***
	(13.99)	(103.54)	(12.64)	(85.39)	(8.37)	(41.38)
(Similarity to grant)^2		-47.391		82.26		-126.546***
		(77.96)		(65.89)		(38.20)
N	1255	1255	4454	4454	1754	1754
R-squared	0.252	0.252	0.125	0.126	0.128	0.136
p-value: joint sig. of sim.	0.983	0.831	0.677	0.212	0.521	0.004**

NOTES: Dependent variable is number of citations received by an article within five years of publication. All models control for # of publications & citations in prior 3 years, age at grant, and publication year fixed effects. Standard errors in parentheses are clustered by award number. *: $p < 0.1$, **: $p < 0.05$, ***: p, < 0.01.

Appendix References

Achakulvisut, Titipat, Daniel E. Acuna, Tulakan Ruangrong, and Konrad Kording. "Science Concierge: A fast content-based recommendation system for scientific publications." *PloS ONE* 11, no. 7 (2016): e0158423.

Aggarwal, C. C., Hinneburg, A., & Keim, D. A. 2001. *On the surprising behavior of distance metrics in high dimensional spaces.* Paper presented at the ICDT.

Beltagy, I., Lo, K., & Cohan, A. (2019). SciBERT: A pretrained language model for scientific text. EMNLP.

Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022 (2003).

Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? Journal of the American Society for Information Science and Technology, 61(12), 2389–2404.

Callaway, B. & Sant'Anna, P. H. C. Difference-in-Differences with multiple time periods. *J. Econom.* **225**, 200–230 (2021).

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. in *Proceedings of NAACL-HLT 2019* 4171–4186 (2019).

Gardner, J. Two-stage differences in differences. 2022. URL https://arxiv.org/abs/2207.05943.

Goldstein, A.P. and Narayanamurti, V. 2018. Simultaneous pursuit of discovery and invention in the US Department of Energy. *Reserch Policy* 47(8): 1505-1512.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. Proceedings of the National Academy of Sciences, 101(suppl_1), 5228-5235, https://www.pnas.org/doi/10.1073/pnas.0307752101.

Hill, R., Y. Yin, C. Stein, X. Wang, D. Wang, and B.F. Jones. "The pivot penalty in research." *Nature* (2025): 1-8.

Lancichinetti, A., Sirer, M. I., Wang, J. X., Acuna, D., Körding, K., & Amaral, L. A. N. 2015. High-reproducibility and high-accuracy method for automated topic classification. *Physical Review X, 5*(1), 011007.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.

Murphy, K. P. 2012. *Machine learning: a probabilistic perspective*: MIT press.

National Academies of Sciences, Engineering, and Medicine. 2017. An Assessment of ARPA-E. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24778.

Office of Energy Efficiency and Renewable Energy. 2015. *2016-2020 Strategic Plan and Implementing Framework*, DOE/EE-2061, https://energy.gov/sites/prod/files/2015/12/f27/EERE_Strategic_Plan_12.16.15.pdf.

Rau, D., Dehghani, M., & Kamps, J. (2024). Revisiting bag of words document representations for efficient ranking with transformers. ACM Transactions on Information Systems, 42(5), 1-27.

Sun, L. & Abraham, S. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *J. Econom.* **225**, 175–199 (2021).

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). ArnetMiner: Extraction and mining of academic social networks. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 990–998). ACM.