Government Funding and the Direction of Academic Energy Research

David Popp[1]
Daniel Acuña[2]
Myriam Grégoire-Zawilski[3]
Lizhen Liang[4]

March 20, 2025

Abstract

Does government research funding influence the topics on which scientists publish? Using data on 699 awards by select Department of Energy and National Science of Foundation programs between 2009-2016, we compare the publication trajectories of funded scientists to a set of matched controls. Our methods combine text and regression analyses to answer 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? For each award, we use topic modeling to identify a control scientist who was not funded by these research programs, but whose research topics were on the same trajectory as the funded researcher, making them just as likely to have pursued the same research directions at the time of the grant. Our control group of scientists 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. 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 be successful. More general calls for funding such as those offered by traditional NSF directorates generate less movement in topic space.

Acknowledgements:
Financial support for this research were 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 Ziqiao Chen, Jingni Zhang, and Hanlin Zhang for excellent research assistance.

[1] Center for Policy Research. Maxwell School of Citizenship and Public Affairs, Syracuse University. 426 Eggers Hall, Syracuse, NY 13244-1020 and NBER. E-mail: dcpopp@syr.edu.
[2] Associate Professor, Department of Computer Science, University of Colorado at Boulder, Daniel.Acuna@colorado.edu
[3] Assistant Professor of Technology and Innovation Policy, La Follette School of Public Affairs, University of Wisconsin – Madison, 1225 Observatory Drive, Madison, WI 53706, gregoirezawi@wisc.edu.
[4] PhD Candidate, Information Science and Technology, Syracuse University, lliang06@syr.edu.

While government-funded energy R&D remains an important part of policy initiatives to address climate change, the political environment for clean energy R&D funding is changing. In the U.S., clean energy faces potential funding cuts and the agencies that support them are under fire. Elsewhere, the responsibility to fund energy research has increasingly shifted from science ministries to energy ministries (Meckling et al., 2022). With greater expertise and capacity to define priorities in the energy field, many of those agencies have shown growing interest in funding use-oriented research that aligns with their mission (Mandt et al., 2020). In this challenging environment, ensuring that research funding is well-spent is vital. To be truly impactful, government investments in clean energy research should not simply substitute for other sources of funding available to researchers but convene scientists around society's important emerging problems that are underfunded. Recognizing this, the Department of Energy's Advanced Research Projects Agency-Energy (ARPA-E) program explicitly lists creating new communities of researchers, including those "not traditionally... involved in the topic area" (National Academies of Science, 2017, p. 122) as a goal.

Yet most academic literature evaluating government research only asks whether funding increases research output, but 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? And 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. Exceptions include Azoulay *et al*. (2011) and Myers (2020), both of which focus on health sciences.

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, we include both mission-driven programs, such as ARPA-E or the DOE's Energy Efficiency and Renewable Energy (EERE) program, and programs featuring open-ended calls for proposals, such as DOE's Office of Science and select National Science Foundation programs. Open-ended programs allow for a broad range of

proposed ideas, typically bounded by a specific discipline (e.g. NSF's Electronic & Photonic Materials) or broad research area (e.g. NSF's Energy for Sustainability), potentially lowering the incentive to switch research topics. We combine data on both calls for proposals and grant abstracts from these programs with publication data of both 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. Our research addresses the following questions:

Do scientists change the focus of their entire research portfolio in response to government funding opportunities? If so, what types of calls for funding best attract new researchers? Are targeted funding opportunities more likely to encourage researchers to change topic, or do they only attract researchers already established in the field? Are open-ended programs in an agency with a single focus (e.g. DOE) more or less effective than similar programs at NSF, which funds scientists in many fields? Finally, some mission-oriented DOE programs use Funding Opportunity Announcements (FOAs) with specific goals for funded research. Researchers can choose the approach they will take, but have less choice over the questions to be addressed. Comparing the text of scientists' research portfolios to the text of the FOA, we can assess both whether these programs move the research of funded scientists closer to the goals of the FOA, or whether these awards fund research on topics likely to be carried out anyway.

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. Johnstone et al., 2010; Verdolini and Gaelotti, 2011; Peters et al., 2012; Dechezleprêtre and Glachant, 2014; and Nesta et al., 2014).

Focusing more directly on the effectiveness of 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. Evaluations of public funding for energy research should take these long-term horizons into account, especially for funding programs in emerging areas of technology whose timelines are inherently uncertain. High-risk/high-reward research grants, such as those allocated under the Department of Energy's ARPA-E program, aspire to achieve long-term research impact, but few studies evaluate their outcomes. Goldstein and Narayanamurti (2018) and a 2017 National Academies of Science report (National Academies of Science, 2017) provide early assessments, but their authors acknowledge that their analyses are only preliminary and not causal in nature. With the initial ARPA-E awards 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. Several recent studies asking whether funding increases the level of research activity find mixed results (e.g. Jacob and Lefgren, 2011 and Rosenbloom et al. 2015). Studies focusing on narrow fields of research however typically find a positive relationship between funding and research outputs (e.g. Blume-Kohout, 2012 on NIH and 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. Our paper differs from these studies by asking does government funding change research directions.

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). Studies show that researchers' work evolves incrementally and that large shifts in research topics over the course of one's career are rare, suggesting that incentives within the scientific profession reward focused research portfolios and the steady deepening of expertise (Jia et al., 2017). However, sudden shocks can compel researchers

to shift their focus onto new topics. Using the COVID-19 pandemic as a natural experiment, Hill and colleagues (2021) 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 in the scientific policy literature investigate the effects of government policies on the direction of research directly. Myers (2016) reviews related work in the field of biomedicine. Much of the work on this topic looks at how changes in stem cell funding by the U.S. government affected researchers. The closest published work to our research is Azoulay et al (2011), who document how the nature of different research grants affects the riskiness of topics chosen. However, they do not consider how funding affects the choice of research field itself. Corredoira et al. (2018) show that federally funded patents have greater influence (e.g. are part of more active and diverse citation trees) and appear more frequently in different technological classes than corporate patents, providing suggestive evidence that federal funds affect the direction of research. However, they cannot identify whether private inventors would have worked in these same areas had there been less federal support. 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. In an event that significantly lowered the cost to access technology, in late 2010 hackers demonstrated how to use Microsoft Kinect video game sensors for other applications, making it an affordable tool for motion sensing research. Furman and Teodoridis (2020) use text-based analysis to show that this lowering of cost of motion sensing research both increased the quantity of research on motion-sensing and encouraged new entrants into the field. Finally, in a recent working paper, Perez-Alanis and colleagues (2024) find that firms that received tax credits to pursue R&D in new directions outside of their core expertise produced more radical innovation. Thus, while there is promising but scarce evidence about how incentives change what researchers do, more systematic studies are needed.

Our research is also 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, Taddy, & Evans, 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 (Zhao et al., 2021, Chai et al., 2019). While some studies use machine learning to devise novel measures of research outputs, none use machine learning to construct a matched sample. For example, Sun et al. (2021) use topic modelling to identify breakthrough inventions: patents whose topics are initially remote, but eventually become mainstream in their field. However, the scope of their paper is to propose a novel measurement approach, not to use this measure in a regression framework.

In their analysis of how the cost of accessing knowledge shapes research directions, Furman and Teodoridis (2020) use machine learning to construct measures of diversity of research profiles. However, they do not apply machine learning to the construction of their matched sample. Other relevant studies that use matching to identify a control group do not leverage machine learning. For example, when investigating whether research published by federally-funded institutions gets cited outside of their field more when it is patented, Shin et al. (2022) find matches for their sample of scientific papers using variables such as the journal's impact factor and the field of research. However, they identify fields using publications' keywords only. By contrast, 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. Similarly, Perez-Alanis et al. (2024) use propensity score to match firms that received R&D incentives to pursue new research directions with firms that did not, using responses from innovation surveys rather than information extracted from the text of these firms' patents. 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 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 consider two dimensions of funding agency strategy: *focus* and *risk-taking*. Focus refers to whether a program accepts proposals on a broad range of basic science (*General*), solicits proposals on a broad-range of clean energy topics (*Open Energy-specific*), or provides funding announcements identifying a single technology along with detailed technical specifications (*Targeted*). *Risk-taking* indicates whether a program was set up specifically to encourage interdisciplinary and/or high-risk/high-reward research—as opposed to other research programs, which we label as "traditional". While all funding programs hope to fund novel, transformative research, traditional research programs don't necessarily encourage people to do research that is different, but rather to continue research they have shown to be good at and that is acclaimed by their peers. Moreover, peer reviewers in these programs may not recognize the merit of innovative ideas they perceive as too risky or too much at the margins of their discipline (Luukkonen, 2012; Boudreau *et al.* 2016). As a result, we hypothesize that both interdisciplinary programs and those with a focus on untested, high-risk ideas will encourage research less similar to a scientist's existing portfolio. Figure 1 summarizes our categorization of the programs included in our study, which we describe in more detail below and in Appendix A.

Figure 1. Categorization of research programs included in our study.

	General	Open Energy-Specific	Targeted Energy
Traditional	DOE Office of Science NSF, Chemical Catalysis NSF, Electronic & Photonic Materials	NSF Energy and Environment NSF Energy for Sustainability NSF Clean Energy Technology	DOE EERE**
Risk-taking	NSF EAGER*	DOE ARPA-E Open DOE ARPA-E IDEAS NSF Sustainable Energy Pathways (interdisciplinary)	DOE ARPA-E focused

*: NSF-EAGER can be either General or Open Energy-specific, since EAGER could be used by any NSF program

**: 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.

We consider three programs at the DOE that span a 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, p. 25). 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 and Narayanamurti 2018), we include additional DOE programs on either side of this spectrum. The DOE's Office of Science Basic Energy Sciences program provides an example of a traditional program supporting basic research using open solicitations. DOE's Office of Energy Efficiency and Renewable Energy (EERE), which supports applied clean energy research, provides an example of a traditional targeted program. 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.

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. 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. Appendix A describes these programs in greater detail. Finally, while supporting transformative research is a goal of all NSF programs, specific NSF initiatives pay particular attention to novel and interdisciplinary research. First, 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 administered in a single directorate, Sustainable Energy Pathways involved five different NSF directorates and required interdisciplinary teams of researchers. Second, all NSF programs consider Early-Concept Grants for Exploratory Research (EAGER) proposals to promote exploratory work on untested, but potentially high risk-high payoff research that "involves radically different approaches, applies new expertise, or engages novel disciplinary or interdisciplinary perspectives" (NSF 2018, p. II-33). We hypothesize that these programs, which emphasize interdisciplinary and/or high-risk, untested ideas, should promote greater changes in research direction than other open-ended programs.

III. Data

The primary data source for this project is the Dimensions database provided by Digital Science. It 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 and 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 (Fig 1), (2) the abstracts of the funded projects, and (3) the proposal solicitations of targeted programs.

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. Finally, to compare researcher records to the initial grant solicitations for our two targeted programs, we obtained the text of funding opportunity announcements from the DOE's Funding Opportunities Exchange website for ARPA-E and EERE.

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 and FOAs across the different programs. In total, we begin with 2,697 grants with starts dates between the years 2009-2016, along with 85 different funding opportunity announcements. 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. Creating a Matched Sample

Our data include the PI's of 2,598 grants awarded by DOE or NSF between 2009 and 2016. We limit the analysis to scientists with US university affiliations in the three years prior to and 5 years after the start date of the grant. 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.

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 matching rule 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 either the funded grant abstract or FOA). 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.

A. Measuring Similarity

To identify matched researchers, we use publication data from the three years prior to the start of the funded research for both the PI and potentially matched researchers. 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. The procedure we describe is agnostic to the sets considered, and therefore enables both comparisons within a scientist or even across a field.

There are multiple ways to compute similarities. We use the text of either the grant's abstract or Funding Opportunity Announcement 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. This is a faster and more accurate way to measure similarities (Achakulvisut et al., 2016; Manning, Raghavan, & Schütze, 2008).

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. Appendix B discusses the calculation of cosine similarity in more detail.

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 either the grant abstract or FOA. 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.

B. Identifying the Control Set

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,[5] and the quality of an author's research institution.[6] Third, we use topic modeling to identify *matched controls*, obtaining at most a single

[5] Calculated based on the date of the author's first publication in Dimensions.
[6] 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. We use the

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. 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. 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. We then apply the following filters to both PIs and their potential matches:

- 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%. The calculation of entropy is described in Appendix B.

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-$pairdist_3$) + $WAdist_3$. 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-$pairdist_3$)+$WAdist_3 \leq$

Carnegie research classification of a researcher's university affiliation to identify researchers in similar environments.

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.

As noted above, 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. 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% (TO BE ADDED).

C. Evaluation of Matching

Our main analysis compares similarity between the grant abstract and each researcher's publications. This sample is larger and allows us to compare the effectiveness of different types of funding programs. We provide additional details on similarity to the FOA in the appendix (Table D3). Table 1 compares the means of key variables for the PIs and controls in our sample. To check for balance in the matched sample, we present the standardized difference of means (SMD) and variance ratio (VR).[7] 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. As shown in the appendix (Tables D1 & D2), while limiting to publication counts within +/- 25% narrows the gap on publication counts, it does so at a cost of dropping

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

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.

Table 1: 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				

Next, we plot similarity to the funded research to compare the evolution of research topics over time for both the PIs and matched controls. In the years prior to funding, both the PIs and their matched controls have similar research profiles (Figure 2). Illustrating the ability of our matching procedure to identify a viable set of controls, Appendix Figure D1 shows almost no relationship between similarity between PIs and the full set of feasible controls. For both PIs and their matched controls, their research profiles become more similar to the grant topic even before funding is granted. Appendix Figures D2 and D3, which plot similarity by calendar year, rather than relative to the start of the grant, help explain this trend. Prior to 2013, the research of both PIs and controls becomes more similar to the grant projects, the bulk of which pertain to clean energy. Similar trends have been documented in other studies of clean energy innovations, including in patents (Popp *et al*. 2022, 2024), venture capital (van den Heuvel and Popp, 2023) and public R&D investment. Clean energy innovation experienced a boom period in the early 2000s, which

peaked 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 2000s boom.

Thus, our data reveal a set of scientists following more general innovation trends by doing more research on clean energy topics in the early 2000s. What is notable is that the trends for PIs and their matched controls diverge after funding. Matching overall trends in clean energy research described above, the research topics of our controls begin moving away from the topics of the grants in our sample, particularly in the larger sample that compares researchers to grant abstracts. In contrast, while PIs do not continue to move closer to their grant topic, funding maintains their attention on related topics longer. Notably, trends in that sample are similar for both NSF and DOE (Appendix Figures D4 & D5), even though not all NSF research topics in our data pertain to clean energy.

Figure 2: Average Similarity to Grant Abstract, years to/from funding

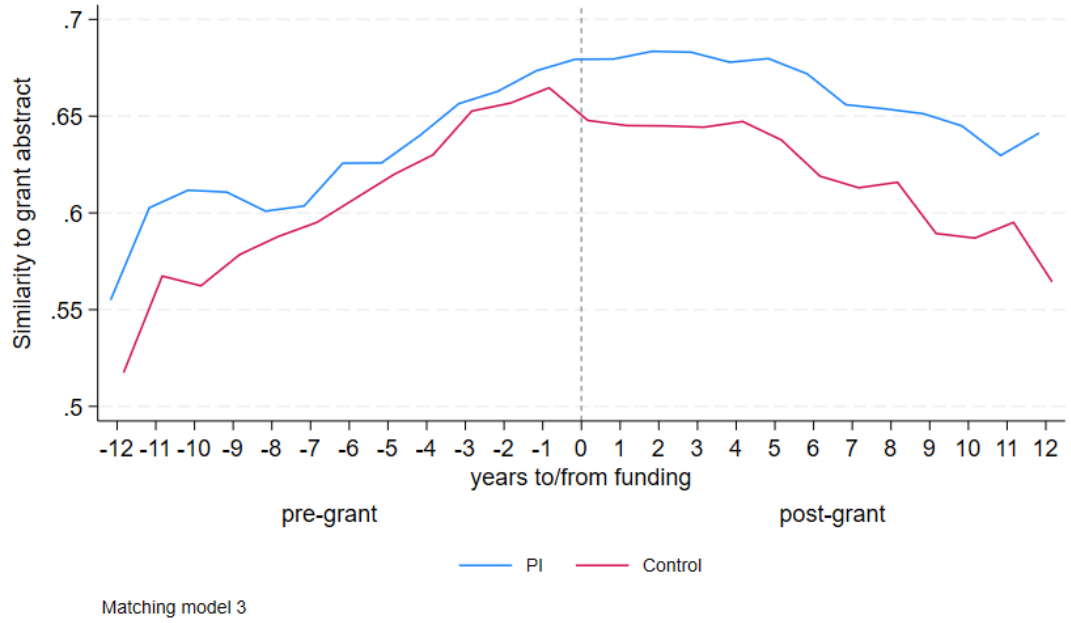
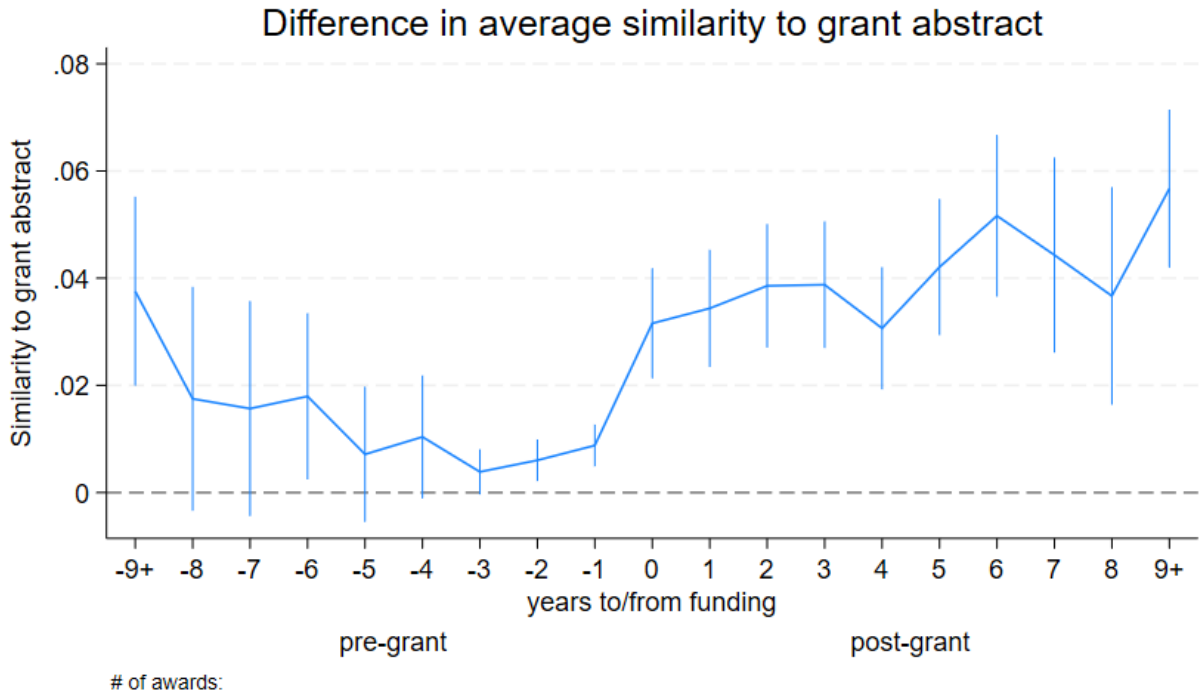


Figure 3 shows the results of a t-test of differences between PIs and their matched controls. In general, both PIs and their matched controls are very similar in the years prior to funding. As suggested by the previous figure, PIs are slightly more similar to the grant than the matched controls. But the differences

are small in the years directly prior to funding with a difference of just 0.0044 in the year before funding, compared to an average similarity to the grant of 0.68. After funding, the research of funded PIs becomes more similar to the grant than the research of their matched controls, ranging between 0.03 and 0.04 more similar in the years directly after funding. The effects persist over time and is drawn from these researchers' entire publication records, suggesting that funding affects not just the publications directly resulting from a grant, but the future research directions of the PI.

Figure 3: T-test of differences between PIs and controls



NOTES: Figure shows the difference in similarity to grant abstract between funded PIs and matched controls. Number of Awards = 699

V. Model Specification

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-1}, and treated in periods $t = t_{FY},\ldots, T$. *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): $E[Y_{i,t}(1) - Y_{i,t}(0) \mid D_{i,t} = 1]$, 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.

However, in Figure 3 we see 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 an event study regression, denoting researchers receiving funding for grant g from funding program p or FOA f 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, we normalize the year prior to treatment equal to 0. Here, since the dependent variable is *similarity to a specific grant abstract or FOA*, 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.

We begin by describing the regression for similarity to a grant abstract. 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. When working with FOAs, FOA-specific year effects allow for different trends over time for applicants (and their controls) to different funding announcements, since different FOAs solicit funding on different topics. Since we do not have funding announcements for the grant abstract analysis, program-cohort-year fixed effects achieve the same purpose of limiting comparisons to those who applied for the same funding competition. 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.[8] Thus, our program and cohort specific year fixed effects ($t_{t,p,FY}$) control for general trends affecting both funded scientists and their controls

[8] 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.

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, Popp (2016) finds that the effect of public funding on publication counts may persist for six years or more. 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 estimates are imprecise at the ends of the sample, when not every researcher is actively publishing. 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,[9] 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*.

Minor adjustments are necessary to apply this model to the analysis using similarity to the FOA, rather than to the grant abstract. For those data, we use the date of the funding announcement rather than the start date of the grant. This avoids complications in our data when different awardees to the same FOA have different starting years. Thus, $FY_{g,t}$ equals one in the due year of the FOA associated with *g* for both the funded scientist and their matched control. Analogous to the program and cohort specific year fixed effects, we include separate year fixed effects for each FOA solicitation.

[9] 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.

As noted above program and cohort specific (or FOA-specific) year fixed effects are important to ensure that the regression only makes valid comparisons across pairs of researchers. As noted by recent literature on two-way fixed effects models (e.g., Callaway and Sant'Anna, 2021; Sun and Abraham, 2020), 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 or FOA. 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, Callaway and Sant'Anna (2021) note that compositional changes in comparison groups can occur in traditional two-way fixed effects event model. 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. [10]

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 do 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

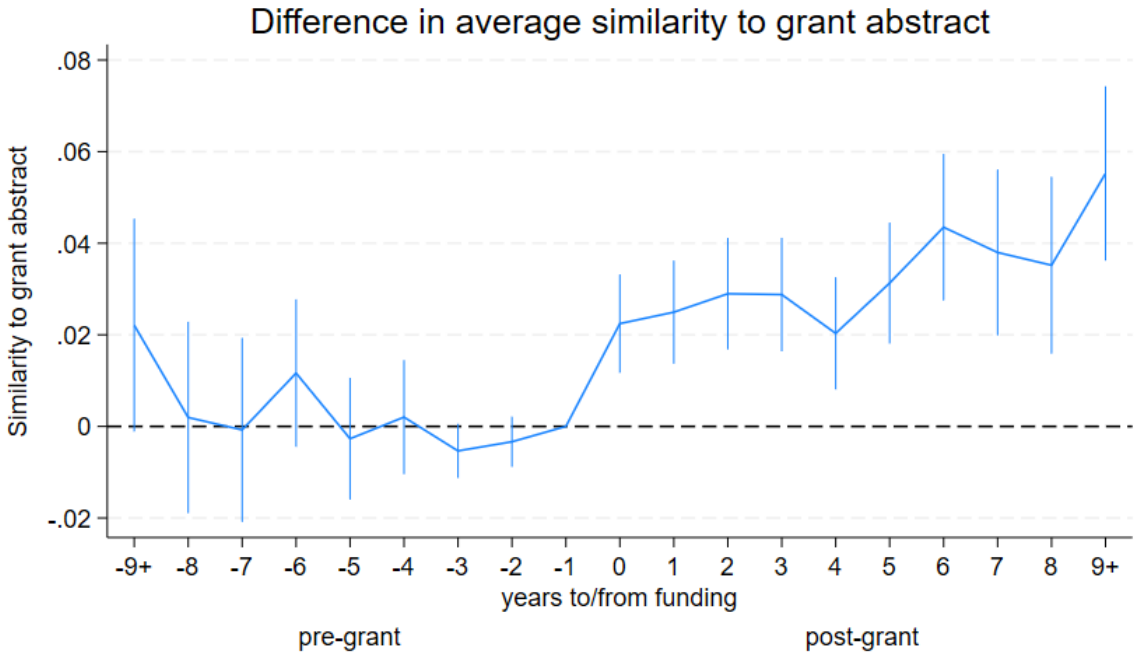
[10] To validate that cohort-specific year fixed effects ensure comparisons are made to the correct controls, note that in a regression replacing individual fixed effects with grant fixed effects and omitting any controls, a regression using cohort-specific year fixed effects exactly replicates the paired t-test results shown in Figure 3. 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.

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

VI. Results

We begin with results comparing similarity to the grant abstract, which allows us to assess whether Targeted programs move research more or less than Open Energy-specific or General calls for funding. Figure 4 presents overall results for similarity to the grant abstract. While the pattern is similar to the simple t-test in Figure 3, notice that the combination of individual fixed effects and control variables remove any remaining small differences between PIs and the matched controls.

Figure 4: Similarity to Grant Abstract, by Program type

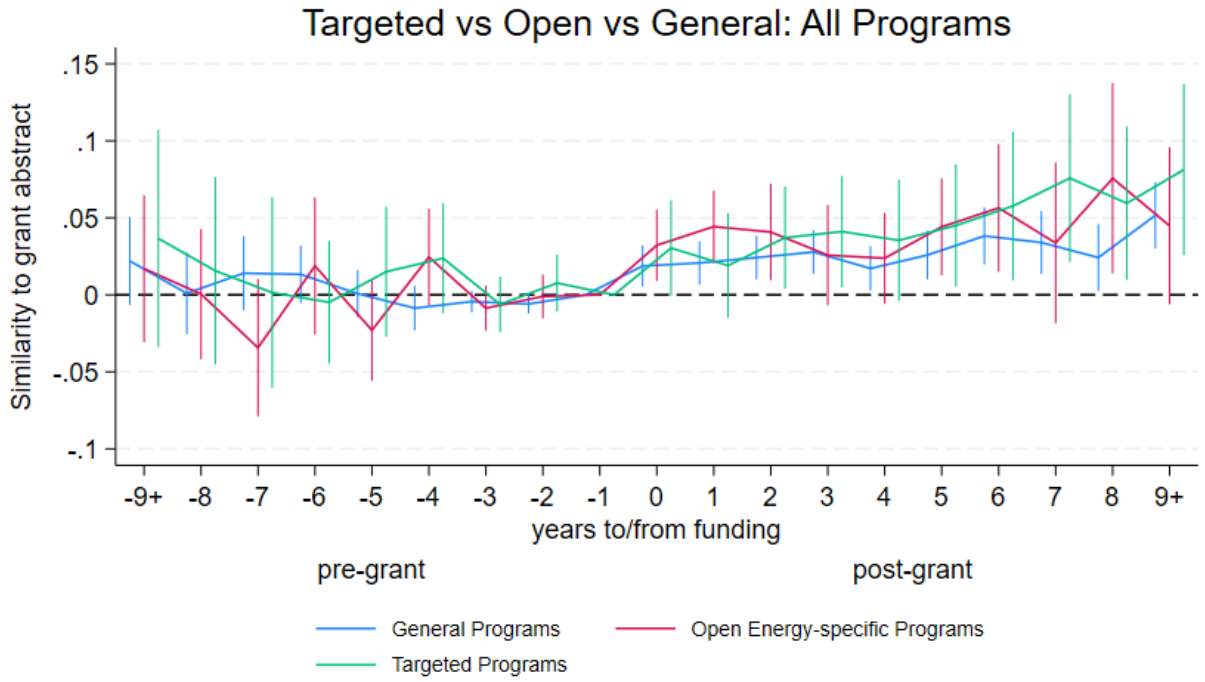


NOTES: All models include individual fixed effects and controls. Number of observations = 24205. Number of Awards = 699.

Figure 5 considers variation across the focus of different programs, pooling both Traditional and

Risk-taking programs (e.g., the columns of Figure 1). Overall, both Open energy-specific and Targeted programs move research closer to the funded topic than General calls for funding. In both cases, the magnitude is roughly twice that of general programs. Notably, the effects for both targeted and open energy-specific programs persist over time, suggesting that these grants help build human capital by keeping researchers in the field. One difference is that open energy-specific programs see faster movement towards the topics than targeted programs.

Figure 5: Similarity to Grant Abstract, by Program type

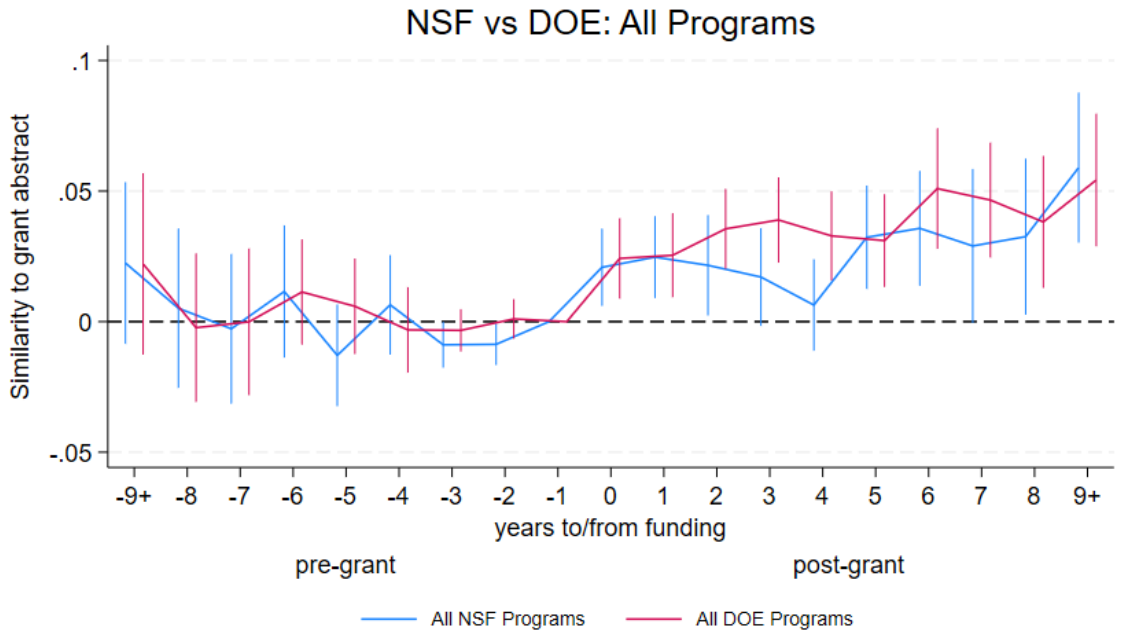


NOTES: All models include individual fixed effects and controls. # of observations: 15968 (General), 4844 (Open Energy-specific), 2710 (Targeted). # of awards: 474 (General), 145 (Open Energy-specific), 80 (Targeted).

To interpret the magnitude of our estimates, recall that the average PI in our sample has roughly 10 publications per year (9 in the abstract sample, 11 in the FOA sample). 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. Using this thought exercise, moving similarity by 0.05 per year, as in the case of Open energy-specific programs, is the equivalent of replacing a paper

half-related to the funded topic to one that is fully related. For General programs, the observed change is about a third of a paper.

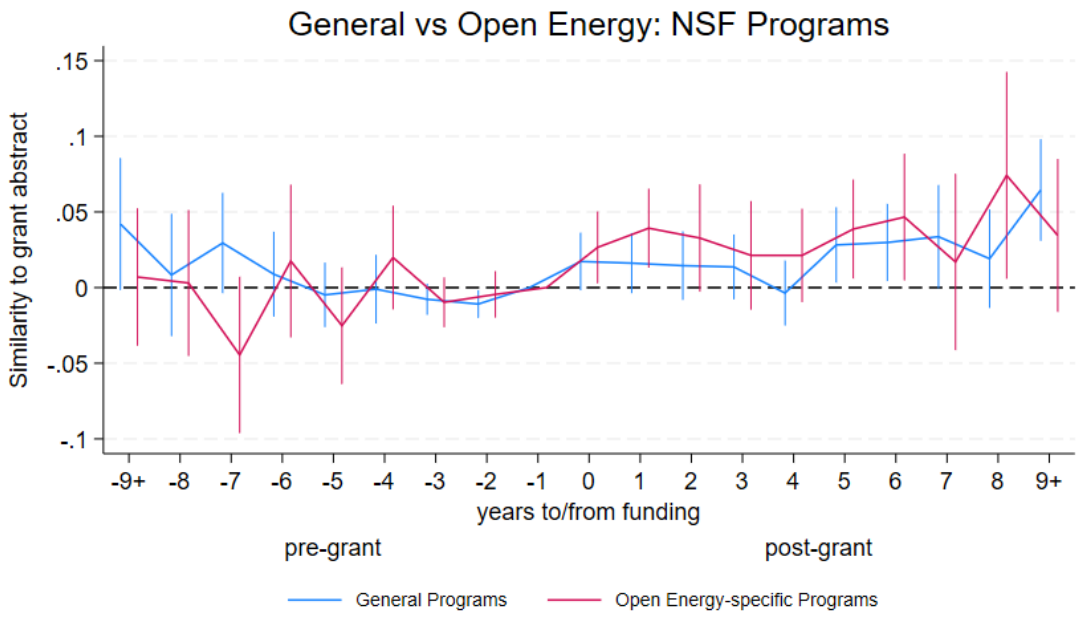
Figure 6: Similarity to Grant Abstract, NSF vs DOE



NOTES: All models include individual fixed effects and controls. # of observations = 11252 (NSF), 12953 (DOE). # of awards = 324 (NSF), 375 (DOE).

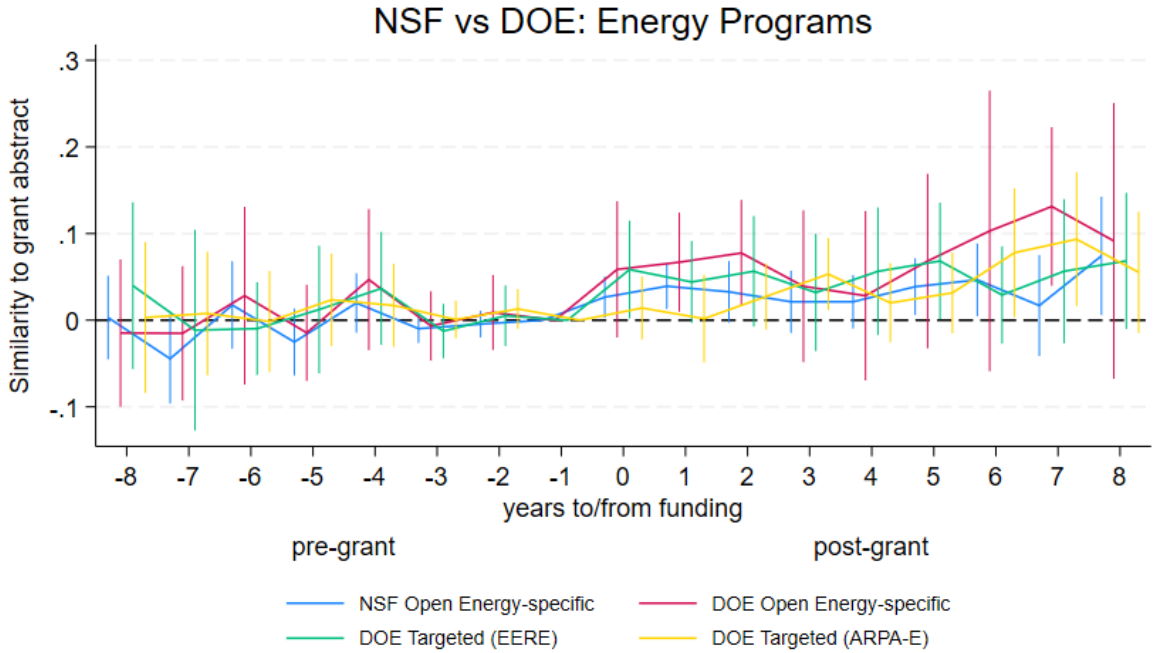
Much of this difference comes from differences between funding from the Department of Energy and the National Science Foundation. In Figure 6, we see that DOE funding leads to a larger and more sustained change in research topics than NSF. Moreover, the slight dip in similarity observed for NSF researchers in the years prior to funding suggests that these researchers may have been moving in the direction of their funded topic even prior to receiving the grant. This difference is partially explained by the larger focus on general funding calls at NSF. In our sample, nearly two-thirds of NSF awards are from General programs. Researchers in these programs may be using NSF funding to move existing research streams forward. Within NSF, Open energy-specific programs produce larger differences in research topics

Figure 7: Similarity to Grant Abstract, NSF Programs



NOTES: All models include individual fixed effects and controls. # of observations: 6930 (General), 4018 (Open Energy-specific). # of awards: 204 (General), 120 (Open Energy-specific).

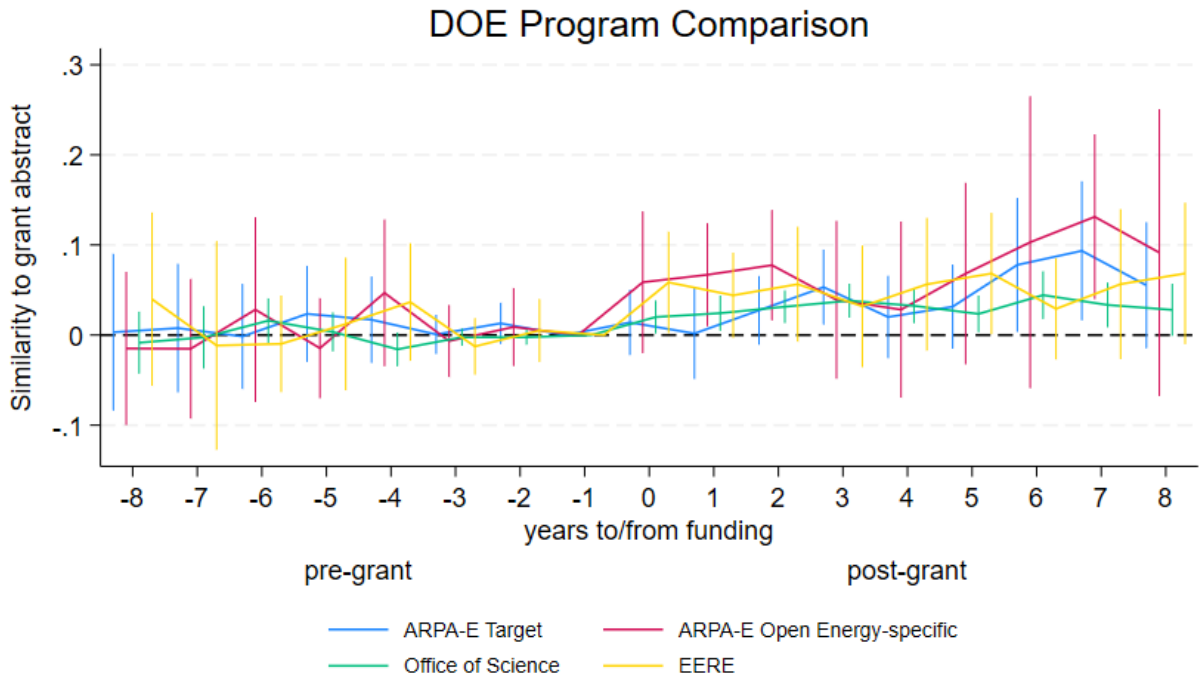
Figure 8: Similarity to Grant Abstract, Energy Programs



NOTES: All models include individual fixed effects and controls. # of observations = 4139 (NSF Open Energy-specific), 821 (DOE Open Energy-specific), 1234 (DOE Targeted: EERE), 1536 (DOE Targeted: ARPA-E). # of awards = 120 (NSF Open Energy-specific), 24 (DOE Open Energy-specific), 36 (DOE Targeted: EERE), 44 (DOE Targeted: ARPA-E). Binned years omitted from figure due to small sample-size at program level.

than General programs (Figure 7). However, the effects are smaller than those of Open ARPA-E and EERE, as well as smaller than Targeted ARPA-E in the long-run (Figure 8).

Figure 9: Similarity to Grant Abstract, DOE Programs



NOTES: All models include individual fixed effects and controls. # of observations: 1504 (ARPA-E Target), 790 (ARPA-E Open Energy-specific), 9038 (Science), 1206 (EERE). # of awards: 44 (ARPA-E Target), 24 (ARPA-E Open Energy-specific), 270 (Science), 36 (EERE). Binned years omitted from figure due to small sample-size at program level.

Within DOE, the key distinction is between high-risk/high-reward programs (ARPA-E) and other programs (Figure 9). ARPA-E's Open awards lead to the largest change in research topics, with magnitudes approaching 0.77 after just two years after funding. 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. Judged by the ability to influence the direction of awardee research, these Open solicitations appear to be a success.

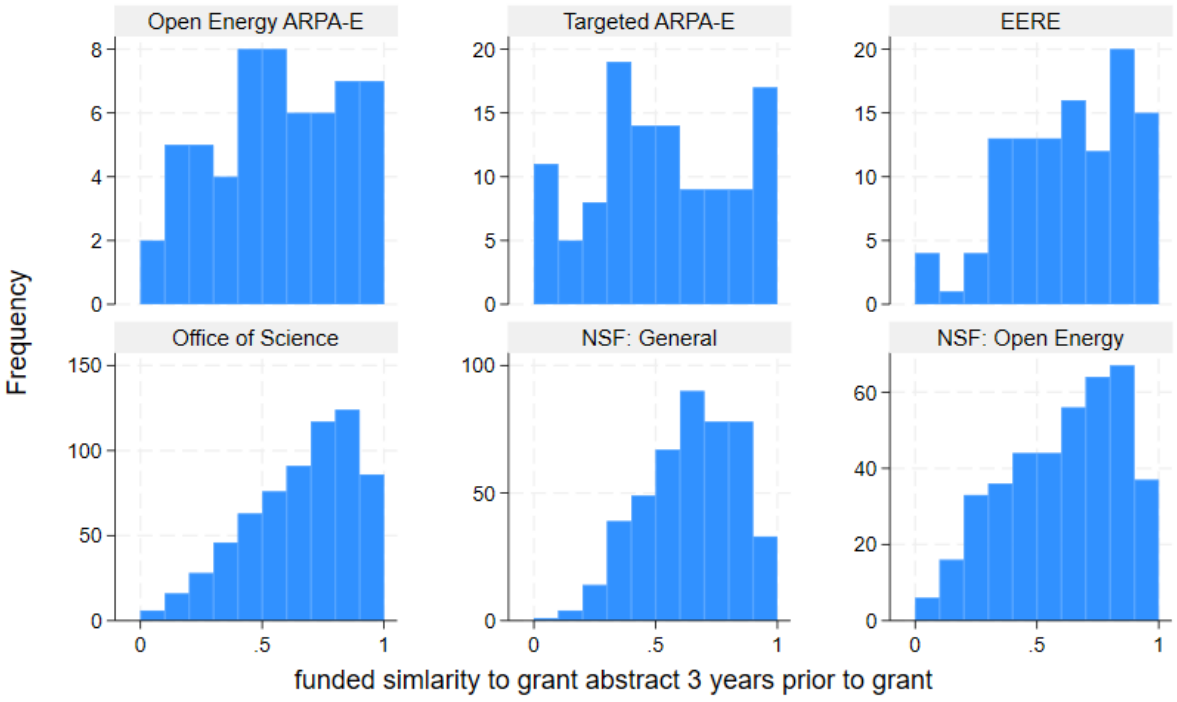
Targeted ARPA-E awards have strong long-run effects, but it takes longer for research topics to

move. That said, given the smaller number of awards when splitting ARPA-E into Open and Targeted awards, the long-run estimates for each are very imprecise. The effect of EERE on topic selection is faster, with an almost immediate effect of 0.56. In contrast, we see further evidence of the more moderate effect of General funding calls, as funding from Office of Science lead to smaller changes in funding, peaking at 0.038 after three years, and again at 0.44 after six years.

That our observed effects are both larger for targeted or 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.

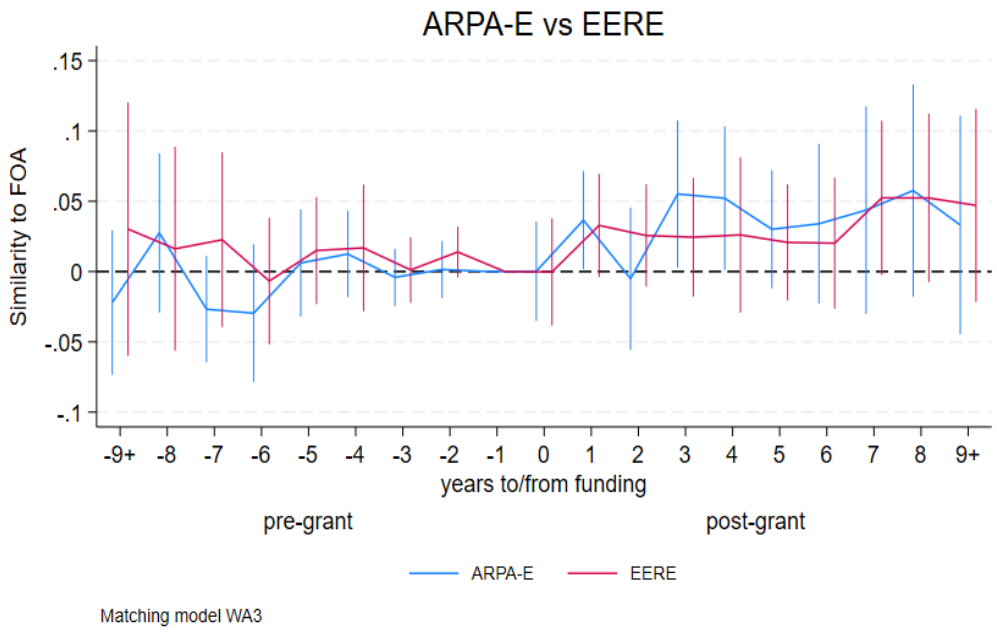
Figure 10 explains why ARPA-E has a larger impact on research direction than other programs. It shows the distribution of the similarity of a PI's research to their grant abstract in the three years prior to funding. In general, the distribution is skewed, with most PIs doing research fairly similar 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 in the graphs below). This result is specific to ARPA-E, as the distribution of similarity to topics at NSF's open energy-specific programs looks similar to more general research funding programs, with the previous research of most PIs being similar to their funded work.

Figure 10: Distribution of similarity to the funded research, 3 years prior to grant



Finally, in addition to whether funding influences the topics scientists choose, our FOA data analysis also allows us to assess whether Targeted funding opportunities move researchers closer to the goals agencies state in their FOAs. Figure 11 presents results for similarity to the FOA. Receiving funding from either EERE or ARPA-E moves the PIs research closer to the FOA of funded project. For EERE, this effect is almost immediate and is steady over time. In contrast, ARPA-E a programs that in addition to soliciting proposals on targeted problems also funds risky projects, leads to a larger change, but that change takes longer to occur. Moving similarity by 0.05 per year, as in the case of ARPA-E, is the equivalent of replacing a paper half-related to the funded topic to one that is fully related. For EERE, the observed change is about a quarter of a paper.

Figure 11: Similarity to FOA



NOTES: ARPA-E # of observations = 1831. # of awards = 53. EERE # of observations = 1601. # of awards = 46. All models include individual fixed effects, FOA-specific year fixed effects, and controls. Matching model 3.

VII. Discussion

Our findings provide 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 ARPA-E open calls. We also find heterogeneity across the different programs. They vary in the duration it takes to achieve full effects and how long these last, with the most persistent effects observed in ARPA-E. Funding programs also vary in the proximity between funded researchers' prior portfolios and the topics of the funded research: the pool of researchers funded by ARPA-E is more diverse in terms of their proximity to the grant abstract. Together these results point to trade-offs that can inform the choice of grant designs. If the goal of funders is to attract researchers new to certain topics, it may take longer to produce outputs closely related to the grant. When researchers have prior experience working on similar topics, they may more readily produce outputs that are closely related to the grant but make smaller pivots. Left unstudied in this paper is whether attracting new researchers to a field produces *better* science. Previous literature

using combination of references to assess novelty suggests that novel works are cited more frequently (Uzzi *et al*., 2013), and that research teams with expertise from more diverse fields and performing multiple research tasks produce more novel research (Lee *et al*., 2015). Whether similar results hold when novelty is measured based on similarity of research topics is left for future work.

Overall, our results show that even if research directions are incremental and path dependent (Jia et al., 2017, Hill *et al*. 2021), they are not immutable. Funders can craft calls that prioritize problems aligned with their missions to compel researchers to shift their focus. Our results consistently show that grants influence the direction of research. Relative to their matched controls, funded scientists do not simply use funding to carry out research that they would have done anyway. Grant programs that seek out 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. Interestingly, the largest shifts occur in the ARPA-E open program – suggesting that researchers shift direction more radically when given space to propose unconventional ideas for solving the problems identified by funding agencies, as opposed to FOAs that call for research in more narrowly-defined technologies to solve these problems. In contrast, conventional grants which also give greater latitude to researchers to choose the problems they focus on without targeting specific problems, yield smaller shifts. We also find that most researchers funded through conventional grants were working on topics related to the grant in the three years leading up to their award. This is consistent with the notion, noted in extent literature (e.g. Luukkonen, 2012), 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. 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 and Kearney 2024).

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 (Rodrik, 2014; Mazzucato, 2018)? 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 from the energy space, our findings have broader generalizability to high-risk/high-reward funding modalities in any domain of science. For example, the Advanced Research Projects Agency for Health (ARPA-H) was launched in 2022 to fund breakthrough health research. While the specific 'missions' defined by the U.S. Department of Health and Human Services ARPA-H program may be different from the ones targeted by ARPA-E, but the ultimate science policy goal of funding transformative research to find solutions to pressing societal problems is the same. Many insights from our paper regarding the success of high/risk-high-rewards funding programs in moving research production toward these pressing problems therefore also apply to other programs, like ARPA-H.

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 other studies (e.g., Popp, 2016) show that grants have a lasting impact - of several years to a decade – on grantees' research programs. In the current 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 2) 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 the clean energy space. 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

to move researchers back into a focus on sustainability and advancing energy transitions.

VIII. Summary and Next Steps

Our methods combine text and regression analyses create a matched set of funded scientists and a control scientist who was similar to the funded scientist in research interest, career stage, and other features up to the time in which our treated scientists start receiving funding. Comparing the evolution of each scientists' research after the treated PI receives funding, we observe that receiving a grant does influence the direction of research. This effect varies across different funding programs. Programs that specifically target clean energy research move research more than programs funding broader science topics. ARPA-E's efforts to attract researchers new to the clean energy field appear successful. ARPA-E attracts researchers interested in a broad range of topics. ARPA-E funding results in a larger move towards the funded grant topic than other funding sources.

At this time, the results presented are preliminary. In the coming months we will add additional robustness checks to verify our results and dig more deeply into descriptive data to better understand the pool of researchers attracted by each program. Among the next steps are:

- Estimating our model for different researcher characteristics. For instance, do the effects vary by experience or by the entropy of the funded scientist? What does this mean for the impact on human capital of any funding cuts?
- Estimate our model dropping papers that acknowledge the grant. This will check whether our results simply pick up papers directly funded by the grant (e.g. researchers are simply doing what they were paid to do) or whether funding builds human capital, leading to long-run changes in research. Even if our results only pick up papers directly funded by the grant, that we observe movement suggests that grants aren't simply paying scientists to work on topics that they would have done anyway. However, evidence that the effect persists would explain macro-level results from Popp (2016), which finds that increases in energy R&D lead to new energy publications even 10 years after funding.

- Add more descriptive statistics on who is funded by different programs. For instance, in our matched sample, ARPA-E attracts scientists who are slightly older and published more papers in the three years prior to the grant.

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.

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.

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.

Boudreau, K.J., Guinan, E., Lakhani, K.R., Riedl, C., 2016. Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance and Resource Allocation in Science. *Management Science* 62(10): 2765-2783. https://doi.org/10.1287/mnsc.2015.2285

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

Corredoira, Rafael A., Brent D. Goldfarb, and Yuan Shi. 2018. Federal funding and the rate and direction of inventive activity. *Research Policy* 47: 1777-1800.

Dechezleprêtre, A. and M. Glachant. 2014. Does Foreign Environmental Policy Influence Domestic Innovation? Evidence from the Wind Industry. *Environmental and Resource Economics*. 58 (3): 391-413.

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

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

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

Hill, R., Yin, Y., Stein, C., Wang, D, and Jones, B.F. 2021. Adaptability and the Pivot Penalty in Science. arXiv preprint arXiv:2017.06476.

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

Jia, Tao, Dashun Wang, and Boleslaw K. Szymanski. 2017. Quantifying patterns of research-interest evolution. *Nature Human Behavior*, vol 1, article #0078.

Johnstone, N. I. Hascic and D. Popp. 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.

Lee, Y.N., Walsh, J.P. & Wang, J. 2015. Creativity in scientific teams: Unpacking novelty and impact. *Research Policy* 49: 684-697.

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.

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.

Myers, Kyle. 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. Washington, DC: The National Academies Press.

Nesta, L., F. Vona, and F. Nicolli. 2014. Environmental Policies, Competition, and Innovation in Renewable Energy. *Journal of Environmental Economics and Management*. 67: 396-411.

Perez-Alaniz, Mauricio, Helena Lenihan, Justin Doran, and Christian Rammer. 2024. Subsidising Innovation Outside or Within Firms' Existing Knowledge Base: Which Is Best for Radical Innovation? *ZEW Discussion Paper 24-007*.

Peters, M., M. Schneider, T. Griesshaber and V.H. Hoffman. 2012. The impact of technology-push and demand-pull policies on technical change–Does the locus of policies matter? *Research Policy*. 41 (8):1296-1308.

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

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): e0138176.

Shin, H., Kim, K., & Kogler, D. F. 2022. Scientific collaboration, research funding, and novelty in scientific knowledge. *PLoS One*, 17(7): e0271678.

Stephan, P.E. 2010. The Economics of Science. Handbook of the Economics of Innovation: vol.1, Bronwyn Hall and Nathan Rosenberg, eds., Academic Press/Elsevier, 217-273.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. 2013. Atypical Combinations and Scientific Impact. *Science* 342(6157: 468-472.

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

Zhao, H., Phung, D., Huynh, V., Jin, Y., Du, L., & Buntine, W. (2021). Topic modelling meets deep neural networks: A survey. arXiv preprint arXiv:2103.00498.

Appendix

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 consider two dimensions of funding agency strategy: *focus* and *risk-taking*. Focus refers to whether a program accepts proposals on a broad range of basic science (*General*), solicits proposals on a broad-range of clean energy topics (*Open Energy-specific*), or provides funding announcements identifying a single technology along with detailed technical specifications (*Targeted*). *Risk-taking* indicates whether a program was set up specifically to encourage interdisciplinary and/or high-risk/high-reward research—as opposed to other research programs, which we label as "traditional". While all funding programs hope to fund novel, transformative research, traditional research programs don't necessarily encourage people to do research that is different, but rather to continue research they have shown to be good at and that is acclaimed by their peers. Moreover, peer reviewers in these programs may not recognize the merit of innovative ideas they perceive as too risky or too much at the margins of their discipline (Luukkonen, 2012; Boudreau *et al.* 2016). As a result, we hypothesize that both interdisciplinary programs and those with a focus on untested, high-risk ideas will encourage research less similar to a scientist's existing portfolio. Figure 1 in the main paper, repeated below, summarizes our categorization of the programs included in our study, which we describe in more detail below.

Figure 1. Categorization of research programs included in our study.

	General	Open Energy-Specific	Targeted Energy
Traditional	DOE Office of Science NSF, Chemical Catalysis NSF, Electronic & Photonic Materials	NSF Energy and Environment NSF Energy for Sustainability NSF Clean Energy Technology	DOE EERE**
Risk-taking	NSF EAGER*	DOE ARPA-E Open DOE ARPA-E IDEAS NSF Sustainable Energy Pathways (interdisciplinary)	DOE ARPA-E focused

*: NSF-EAGER can be either General or Open Energy-specific, since EAGER could be used by any NSF program

**: 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.

We consider three programs at the DOE that span a 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, 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.[11] The Basic Energy Sciences program "supports basic scientific research to provide foundations to lay the foundations for new energy technologies and to advance DOE missions in energy, environment, and national security."[12] 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

[11] 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.

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

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). EERE provides an example of targeted programs not classified as high-risk/high-reward (i.e., traditional).

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.

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. As the audience of potential grantees at NSF is likely larger than DOE, understanding the impact of such programs at NSF is important. We include three "traditional" NSF programs with an energy-specific focus. 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.[13] 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, we include awards made in the two general programs listed above in response to additional funding made

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

available across multiple directorates for Clean Energy Technology. In select years, NSF made additional funds available for clean energy across multiple directorates. See, for example, a 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 disciplinary area. These are identified in NSF's funding database with reference code 8396.

Finally, while supporting transformative research is a goal of all NSF programs, specific NSF initiatives pay particular attention to novel and interdisciplinary research. First, 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. Second, all NSF programs consider Early-Concept Grants for Exploratory Research (EAGER) proposals to promote exploratory work on untested, but potentially high risk-high payoff research that "involves radically different approaches, applies new expertise, or engages novel disciplinary or interdisciplinary perspectives" (NSF 2018, p. II-33). We hypothesize that these programs, which emphasize interdisciplinary and/or high-risk, untested ideas, should promote greater changes in research direction than other open-ended programs.

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 or FOA 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 or Funding Opportunity Announcement and the abstract of each researcher's publications as the primary data to estimate similarities. Therefore, we use Natural Language Processing (NLP) techniques to find *semantic* similarity between sets of documents based on their raw text. This is a faster and more accurate way to measure similarities than network-based similarity measures (Achakulvisut et al., 2016; Manning, Raghavan, & Schütze, 2008).

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:

$$c = \frac{1}{L}\sum t_z \tag{3}$$

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 724,834 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, grant abstracts, and FOAs.

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.

Measuring Similarity and Breadth of Research. With topic distributions computed for both research papers and FOAs, 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 grants or FOAs, we have two set 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:

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

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 ($WAdist_3$).

Step 1) Create the universe of potential controls: 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 724,834 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

control. For the FOA analysis, we only remove PIs of other grants that have an FOA, and are thus included in that regression.

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 39,069 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 scientists 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 either the funded grant abstract or FOA). 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-$pairdist_3$) + $WAdist_3$. To ensure that the match is a quality match, we only keep potential matches where (1-$pairdist_3$)+$WAdist_3 \leq 0.15$ 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 ….BRIEFLY DESCRIBE VALIDATION HERE.

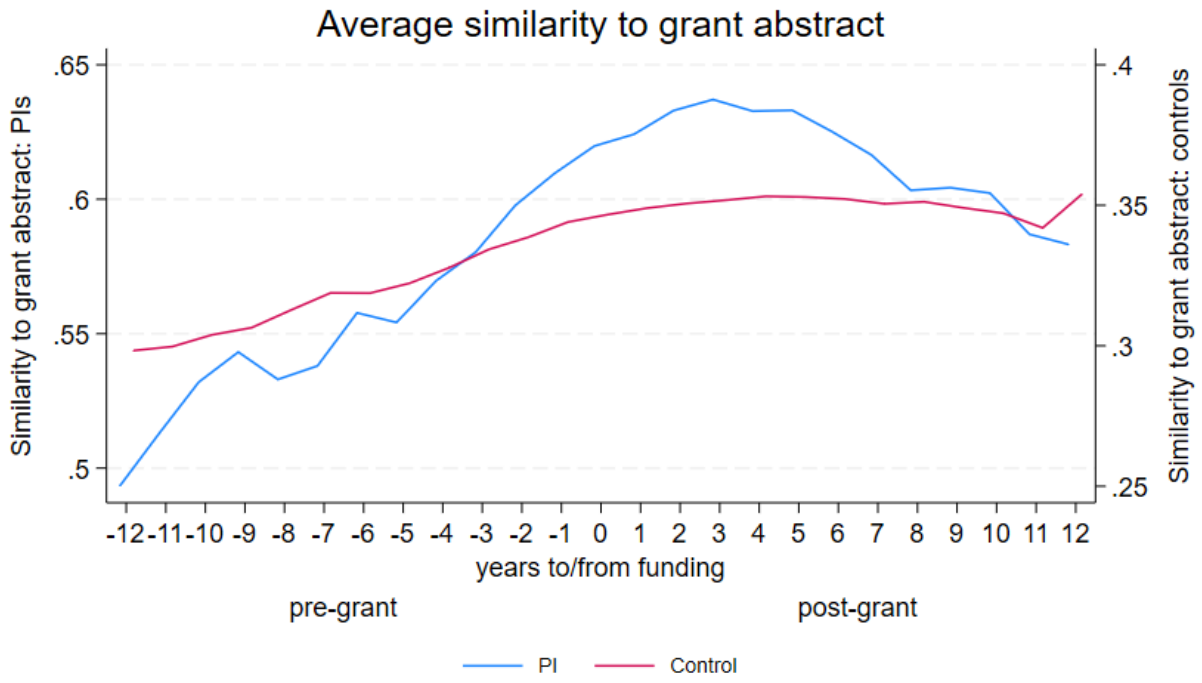
We consider different matching rules to check the validity of our assumptions. Each alternate matching rule differs based on how similar the *level* and *quality* of each researcher's publication record must be. For the *level* of activity, we restrict the number of publications in the 3 years prior 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 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

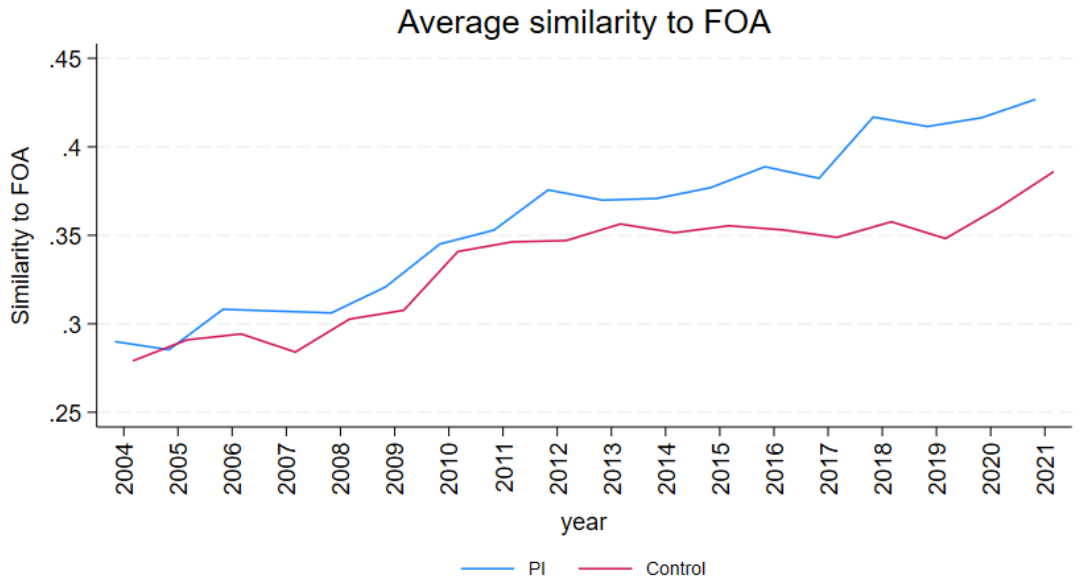
other observable characteristics such as age and research institution. As we see, relative to our PIs, the research of *potential* controls is nearly half as similar to the grant topic and moves little over time, fluctuating between 0.3 and 0.35. Once we identify matched controls by also considering 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.

Appendix Figure D1 Similarity to Grant for PIs and All Feasible Controls

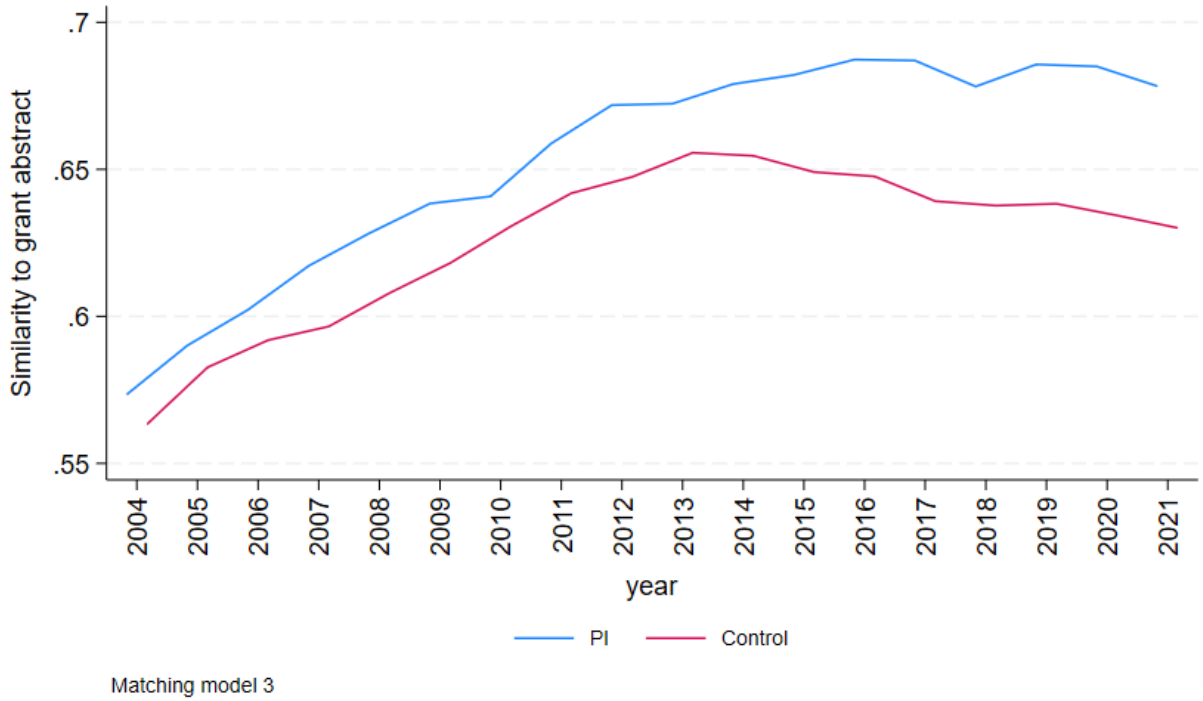


NOTES: Plots the average similarity to the grant abstract for PIs and all feasible controls. To facilitate comparisons of trends across the two groups, the left and right axes have the same scale, but a different range.

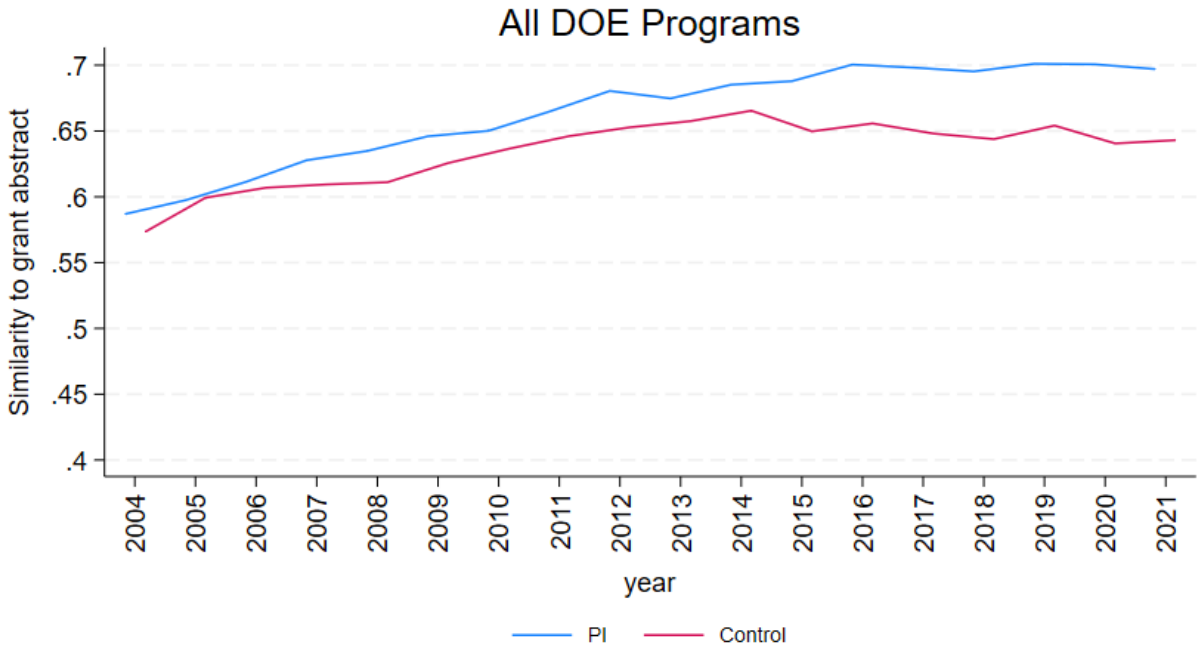
Appendix Figure D2: Average Similarity to FOA, by year



Appendix Figure D3: Average Similarity to Grant Abstract, by year

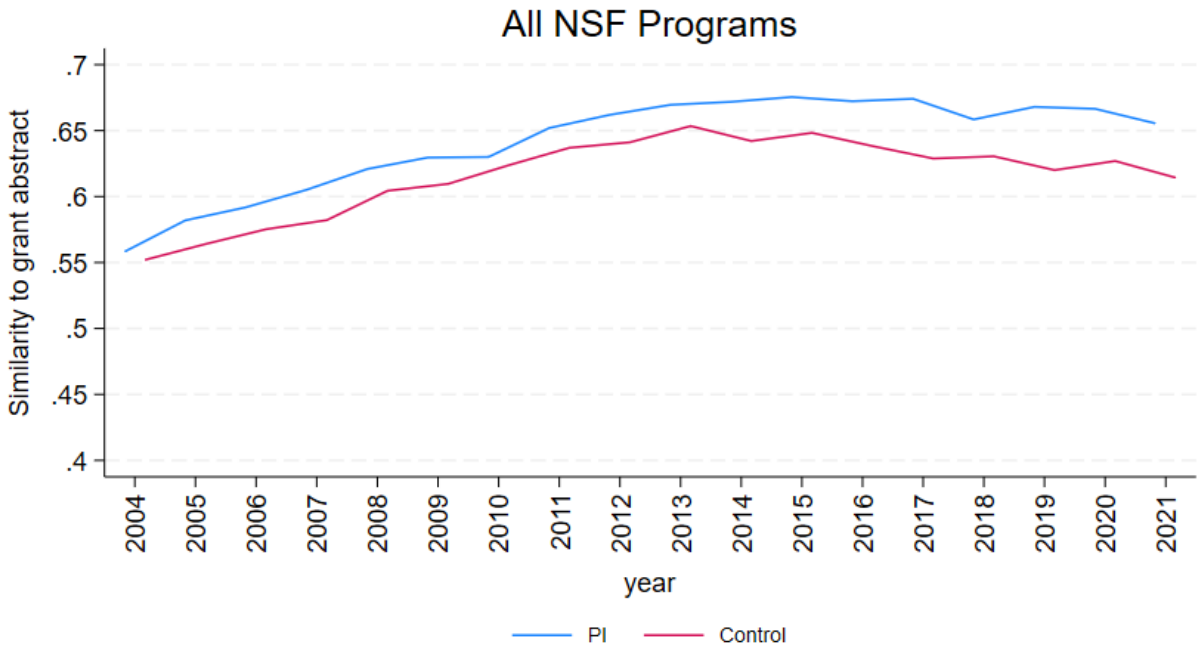


Appendix Figure D4: Average Similarity to Grant Abstract, DOE



Matching model 3

Appendix Figure D5: Average Similarity to Grant Abstract, NSF



Matching model 3

Other Matching Rules

Table D1: 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 D1: 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				

*** RESULTS WITH DIFFERENT MATCHING RULES TO BE ADDED ***

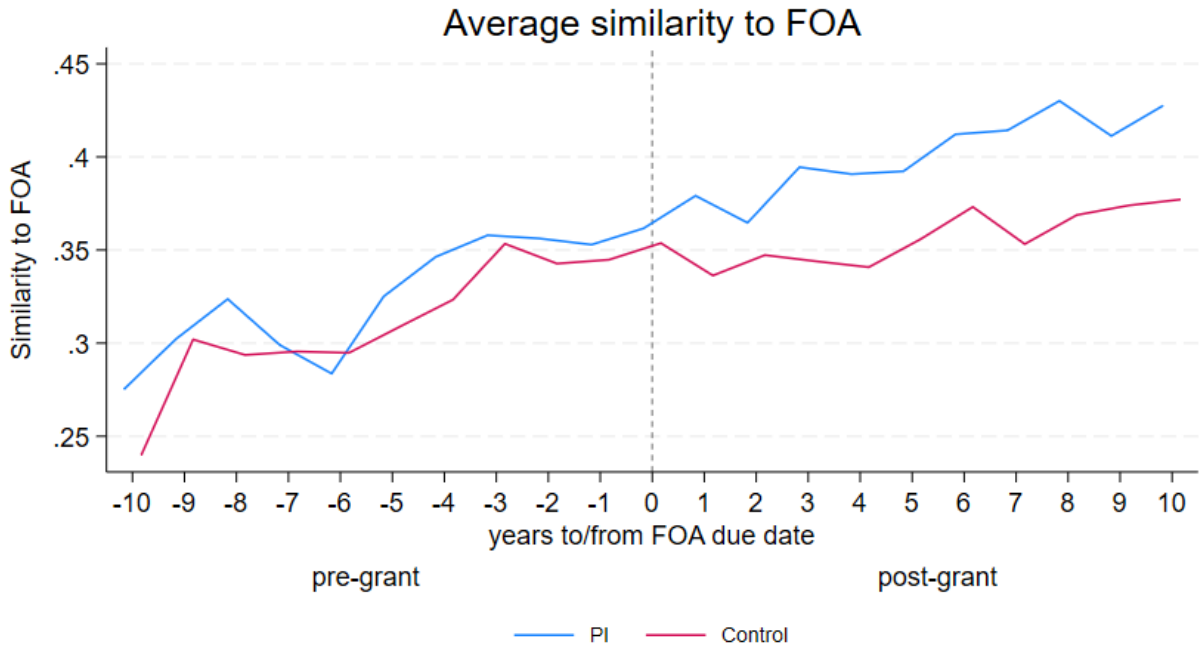
Similarity to the FOA

Turing to similarity to the FOA, we see close matches on most variables. However, the smaller sample size results in some significant differences. PIs have nearly three more years experience (SMD = 0.292) and more co-authors (SMD = 0.204, but VR = 8.558). However, the discrepancy on co-authors comes from a single PI who published one paper with over 1000 co-authors. As shown in the appendix, by allowing for more feasible matches, not considering citations narrows the experience gap, but results in a larger than acceptable gap for the number of publications (SMD = 0.266).

Appendix Table D3: Quality of Matching, FOA Analysis (N_{pub} +/- 50% &(N_{cites} +/- 50%)

	Mean (treated)	***Mean (control)***	***SMD***	***VR***
Similarity to FOA (3 yr prior)	0.363	0.353	0.033	0.976
	(0.291)	(0.295)		
Experience at grant	23.548	20.852	**0.292**	1.494
	(10.536)	(9.346)		
Average citations (3 yr prior)	53.616	51.789	0.043	1.258
	(45.282)	(40.367)		
Citations (3 yr prior)	1956.151	1731.343	0.084	1.234
	(2801.688)	(2522.260)		
N co-authors (3 yr prior)	69.465	53.192	0.204	**8.558**
	(106.530)	(36.415)		
Entropy (3 yr prior)	1.920	1.865	0.153	0.927
	(0.357)	(0.371)		
N publications (3 yr prior)	33.131	29.919	0.158	1.339
	(21.797)	(18.836)		
# of awards: 99				

Appendix Figure D6: Average Similarity to FOA, years to/from funding



Matching rule 3. # of awards: 99

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.

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.

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.

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.

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.

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.