American Economic Journal: Applied Economics 2020, 12(4): 103–134
https://doi.org/10.1257/app.20180518

The Elasticity of Science[†]

By Kyle Myers*

This paper identifies the degree to which scientists are willing to change the direction of their work in exchange for resources. Data from the National Institutes of Health are used to estimate how scientists respond to targeted funding opportunities. Inducing a scientist to change their direction by a small amount—to work on marginally different topics—requires a substantial amount of funding in expectation. The switching costs of science are large. The productivity of grants is also estimated, and it appears the additional costs of targeted research may be more than offset by more productive scientists pursuing these grants. (*JEL* H51, I10, I23, O31, O33)

The efficiency of any market hinges on the ability of its actors to redirect their resources to new opportunities. In the market for science, these opportunities may come in the form of technological breakthroughs, demand for new knowledge, or, as is the focus of this paper, government intervention. Given that scientists are a key source of ideas that drive economic growth (Stephan 1996), it is important to know how costly it is to incentivize changes in the direction of their work—the elasticity of science. This paper provides the first estimates of these costs based on a novel administrative dataset of targeted funding at the world's largest scientific agency, the US National Institutes of Health (NIH).

While the rationale for the public support of science has been appreciated (see Nelson 1959, Arrow 1962), it is often assumed that when governments direct funds to certain pursuits, scientists will simply follow. But how costly, for example, would it be to incentivize a scientist studying one disease to pursue another? And how would this redirection affect their productivity? Quantifying these parameters is necessary for predicting the value of research policies. But empirical challenges have limited progress despite it being more than 50 years since economists set their sights on understanding the rate *and direction* of inventive activity (NBER 1962). Identification issues abound since new opportunities arise endogenously and

* Harvard Business School, Soldiers Field, Boston, MA 02163 (email: kmyers@hbs.edu). Seema Jayachandran was coeditor for this article. I am grateful to Matt Grennan, David Hsu, Ashley Swanson, and Dan Levinthal for their insights and support in developing this paper as a part of my dissertation. I also received many helpful comments from Ashish Arora, Pierre Azoulay, Kevin Boudreau, Amitabh Chandra, Barton Hamilton, Josh Krieger, Mark Pauly, David Popp, Bhaven Sampat, Scott Stern, Heidi Williams, and many seminar participants. This project would not have been possible without Misty Heggeness as well as the rest of the staff of NIH's Office of Extramural Research. I am indebted to W. John Wilbur for taking the time to reconfigure his software. The views and opinions expressed in this paper do not necessarily reflect those of the NIH or any other federal agencies. Funding from the Kauffman Foundation and the NBER program on the Value of Medical Research is gratefully acknowledged. All mistakes are my own.

[†] Go to https://doi.org/10.1257/app.20180518 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

scientists self-select their pursuits. Accounting for these factors is compounded by the difficulty of quantifying a scientist's "direction."

This paper overcomes these challenges by examining targeted grant opportunities at the NIH—which request particular types of science—and a validated algorithm that measures the "scientific similarity" between two abstracts—i.e., how similar a scientist's publications are to the science requested by the NIH. By evaluating scientists' decisions to pursue certain opportunities as a function of the amount of funds made available and how similar the opportunity is to their expertise, I can shed light on the elasticity of scientists' direction.

The NIH awards the majority of grants through so-called investigator-initiated competitions, which cater to all types of biomedical research. But routinely, the NIH sets aside funds for one-time competitions, which request proposals on specific diseases, populations, and/or methodologies. These are termed Requests For Applications (RFAs) and usually look to award $2 to $3 million. Consider this remark by NIH Director Francis Collins from a March 7, 2016 NIH ME/CFS Advocacy Call :

> [The NIH] is working to define the strategic areas of research that would form the basis for a request for applications [RFA] … We're quite serious about looking for opportunities to expand our research in this area and to recruit new investigators into the field, bringing new eyes and new brains into the issue.[1]

Clearly, the NIH believes it can steer researchers to certain topics. However, there is no evidence on how costly it is to bring "new brains" into a field or whether these new brains can contribute.[2] To get an initial sense of the importance of scientific similarity and fund availability, see Figure 1. Figure 1, panel A shows that scientists are much more likely to enter an RFA when the research objectives of that RFA are more similar to their prior work. Figure 1, panel B shows that scientists prefer to enter RFAs with more funds available. Both results are intuitive, but this is the first clear illustration of these facts.

With these data in hand, I first show that it is possible to induce scientists to shift their research focus, but incentivizing these redirections requires a substantial amount of funds. By revealed preference, it appears these sort of adjustment costs are very large. I then show that in equilibrium, even given the large number of scientists in play, RFAs must make more funds available to attract the same number of applications as the investigator-initiated mechanism. Finally, I compare the productivity of grants awarded via these two mechanisms to shed light on the net costs and benefits.

This setting is well suited, as it allows me to quantify typically unobservable variables (i.e., redirections and payoffs) for a large, diverse set of real-world competitions and, essentially, the universe of potential competitors. Given the large costs

[1] Source: https://goo.gl/32Dmr9, accessed July 12, 2017.

[2] Economists have appreciated the adjustment costs of changing input levels (cf. Oi 1962); but only recently has work identified directional adjustment costs of research and development (R&D) activities, most of which has focused on private firms in the energy sector (i.e., Aghion et al. 2016).

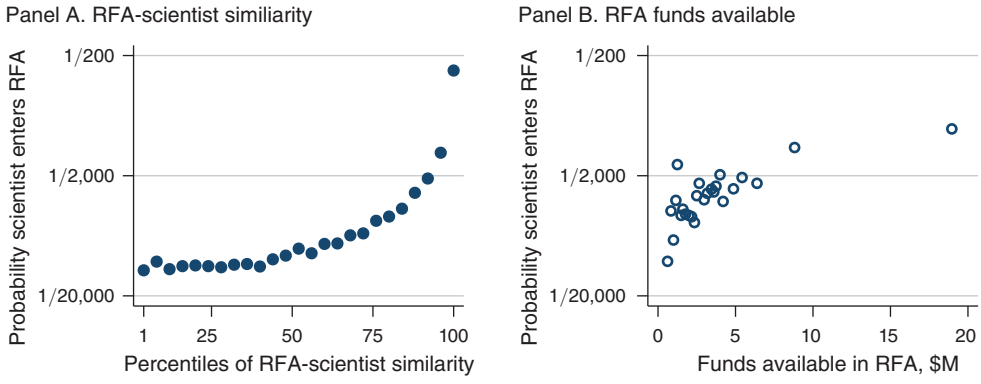

FIGURE 1. PROBABILITY OF RFA ENTRY PER SIMILARITY AND FUNDING

Notes: The figure shows binned scatterplots of entry probabilities per panel A, similarity of scientists' prior publications to the research objectives of the RFA (larger scores indicate greater overlap), and panel B, the amount of funds made available in the RFA. The figure is based on approximately 110,000 scientists and 390 RFAs. Note the log scale of the *y*-axis.

of modern science, this is a useful, practical alternative to a randomized experiment. Furthermore, it allows me to test whether the funds directed by the NIH are actually used as intended.[3]

My results complement a long line of sociological studies that emphasize the role of nonpecuniary forces in science (i.e., Crane 1969, Kuhn 1970, Merton 1973). More recently, Stern (2004) identifies a large wage premium associated with the right to publish, indicating scientists have strong preferences over their broad career direction. Closely related is Azoulay, Fons-Rosen, and Graff Zivin (2019), which also studies how scientists choose which topics pursue. They find that, following the unexpected deaths of preeminent scientists, individuals from outside fields are more likely to enter and succeed; superstars create barriers to entry.[4] I build on this work in identifying how pecuniary incentives can shape a scientist's direction given the cumulative role of preferences, perceived barriers, and any other constraint. The costs I identify do not directly imply any frictions. But it is rather surprising that virtually zero funds from any major scientific funding agency in the world explicitly subsidize field-to-field transitions.[5]

The paper proceeds as follows. First, Section I describes the data and policy variation used, paying attention to the potential endogeneity of RFAs. In Section II, I focus on scientists' decisions to enter RFAs depending on the funds made available and how similar the RFA is to their prior work. The analyses illustrate which scientists are most responsive to RFAs and allow me to infer that there are large

[3] Recall that Nelson (1959) and Arrow's (1962) conclusion is not to simply fund "science" generally speaking but to subsidize the *specific* activities with the largest wedge between the marginal social and private returns. In other words, policymakers need an effective mechanism for selectively targeting funds, such as RFAs.

[4] See also Moser, Voena, and Waldinger (2014) and Borjas and Doran (2015), which examine similar shocks to inventors and mathematicians.

[5] To my knowledge, such an award does not exist for the National Science Foundation, the European Commission, or the Medical Research Council. The NIH has one grant mechanism with related objectives, the "K18," but these grants prespecify the destination field and are rarely used.

directional adjustment costs facing scientists more generally. These costs are estimated using an econometric approach to handle the endogeneity of competition that might bias reduced-form estimates.[6]

Given the large adjustment costs identified, Section III evaluates to what extent they give rise to any "RFA premium." Since it is clear that the adjustments needed to enter RFAs are costly for most scientists, this may require RFAs to use more funds to attract researchers than the alternative investigator-initiated grants (where scientists choose their own direction). However, the pool of scientists that consider NIH grants is large and diverse, so the equilibrium outcome is unclear. I find that RFAs do in fact require more funds to attract new applications, and the size of this premium is about what would be expected given the elasticities identified in Section II.

Then, in Section IV, I use an instrumental variable (IV) to estimate the productivity of RFA and investigator-initiated grants. The results indicate that RFAs are more productive on a per-grant and per-dollar basis. But this appears to be driven by the different composition of scientists and projects, as the productivity gap can be eliminated with a comprehensive set of covariates. Focusing on RFAs, I find that they do induce scientists to move their work closer to the objectives of the RFA, but it appears to be only temporary. Section V concludes.

I. Setting and Data

A. *NIH Overview*

Broadly speaking, the NIH's priority is to award roughly $28 billion each year in grants to scientists based at universities, medical centers, and other research institutions. The key mechanism through which the NIH attempts to steer these funds, and thus the direction of science, is RFAs. Including all major types of research grants, RFA awards have grown as a share of the budget from less than 5 percent in the 1980s to roughly 30 percent as of 2015.

To clarify the role of the NIH in the scientific funding landscape, it is, by a large margin, most scientists' preferred funding source. NIH grants are commonly viewed as a signal of quality (e.g., awards are displayed on individuals' CVs). Scientists take great care to stay abreast of the NIH and commit extensive time to grant pursuits. Application decisions are not taken lightly.

The grants I study provide funds for "projects." Applications for these grants propose a self-contained research idea. Funds can be used to buy inputs (e.g., equipment, materials), pay for travel, or subsidize salaries.[7] I examine the most common research grant type awarded, the R01, which accounts for about 60 percent of all grant awards (70 percent of funds). These are "award[s] made to support a

[6] If strategic interactions are unaccounted for, then I would underestimate scientists' valuation of these funds, since with each funding increase comes additional competition and, therefore, a reduced chance of winning those funds. Thus, I must effectively hold each scientist's chance of winning fixed.

[7] Awards have two components: direct and indirect costs. Direct costs depend on the specifics of each project and are managed at the discretion of the scientist. Indirect costs are based on institution-NIH negotiated rates to support overhead. Because indirect costs reflect institutional differences, I focus on direct costs. Robustness tests in online Appendix F show the main results hold when examining total costs.

discrete, specified, circumscribed project." In my sample, the average R01 awards roughly $285,000 in the first year and lasts for 4.2 years.

When seeking funding at the NIH, scientists have two major options: RFAs and the "investigator-initiated"—or, for brevity, what I will refer to as "open"—grants. Online Appendix A outlines the application process in detail. In brief, all applications are submitted, peer reviewed by panels with similar expertise, and sorted by review scores for funding priority. These processes occur separately for RFA and open applications and at much different scales and timing. In the open mechanism, scientists submit proposals that compete in very large, recurring competitions. In contrast, in an RFA, funds from one or more NIH Institutes are set aside for a single, one-time competition related to a predefined area of science.[8]

The key differences between RFA and open grant competitions are as follows. Each open competition is roughly 20 times larger than an RFA in terms of the total funds and eventual awards. While one peer review panel is convened specifically for each RFA, applications in an open competition include submissions from 65 different peer review panels on average (drawn from the roughly 175 standing peer review panels); the breadth of science is much larger in open competitions. Open applications have a win probability of 16.3 percent with an average size of $275,000, whereas RFA applications have a win probability of 19.4 percent with an average size of $339,000. First-time RFA applications are both more likely to win and, conditional on winning, are larger. Section III focuses on this apparent wedge to identify how much of it is due to the costs facing scientists when adjusting the direction of their work.

B. *Data Sources*

NIH Applications and Awards.—Data on all grant applications to the NIH from fiscal years 2002 to 2009 were obtained from the NIH's administrative database. The full data contain the following: application and applicant identifiers, peer-review grouping and score, funding decision and award size, and institute and fiscal year. For applications submitted in or after 2006, the data also contain the abstract and title of the proposed research project for both funded and nonfunded applications. Only "new" applications are included in all analyses. These are proposals that are being submitted to the NIH for the first time and are not directly tied to any ongoing NIH projects the investigators may have at the time of application.[9]

RFA Details.—The research objectives, funds allotted, and timing of each RFA announced between fiscal years 2002 and 2009 are scraped from the NIH announcement website and manually reviewed to ensure accuracy.[10] A total of 1,125 RFAs are scraped. I restrict the sample of RFAs to include only those that solicit R01 grants

[8] In both cases, there are 20–25 unique competitions of either type available at any given time, although it is always the same group of open competitions, which are organized according to the NIH's institutes.

[9] The NIH has complicated processes for resubmissions of rejected applications as well as the extension of funding for ongoing grants. I was not able to access data on these sorts of applications. Regardless, the new applications I focus on are clearly a policy-relevant subset.

[10] Available at https://goo.gl/LuaBOQ, accessed July 12, 2017.

(686), were not released as a part of the American Reinvestment and Recovery Act (678), and do not request "renewal" grant applications (537).[11] In the entry model, I focus on 394 RFAs, excluding those that do not explicitly state the amount of funds expected to be allotted. In the premium analysis, I examine 453 RFAs released between 2006 and 2009 because this approach requires the applications' abstract data, which are only available in applications 2006 onward.[12]

Publications.—Each scientist's publication record (regardless of funding) prior to 2009 is constructed using the disambiguated version of the PubMed scientific article database developed by Torvik and Smalheiser (2009). PubMed is the National Library of Medicine's database of publications and is considered the gold standard library of biomedical research. Torvik and Smalheiser (2009) develops an algorithm for computing clusters of articles that belong to the same inferred author, which has shown to be extremely precise for NIH-funded scientists in particular (Lerchenmueller and Sorenson 2016).

C. *RFA Generation and Endogeneity Tests*

It is important to clarify how RFAs are generated and the extent to which the objective function of NIH staff should influence the interpretation of the results. In short, RFAs are generated in response to political forces, NIH's programmatic preferences, and other events such as budget shocks.[13] This is relevant to two aspects of this study: (i) identification: are scientists responding to the NIH's funds or other correlated events?; and (ii) generalizability: how representative are scientific topics targeted by RFAs of the full spectrum of science?

The question then is to what extent do these motivations select particular types of science at particular points in time? to test for any kind of differences between the "types" treated by RFAs, I must first discretely classify scientific topics. Thankfully, the National Library of Medicine (NLM) maintains a comprehensive dictionary of scientific terms called Medical Subject Headings (MeSH). The NLM systematically assigns a set of relevant MeSH terms to every publication in PubMed and has made the natural language processing tool underlying this process publicly available. Using this tool, I generate a panel dataset that describes, in each time period, how many PubMed publications and NIH applications are related to a particular MeSH term as well as whether or not an RFA targeted that MeSH term in that period.[14]

Online Appendix B provides further details about sample construction and displays the cross-sectional distribution of treated and control terms. In the cross section, RFA-treated terms occur in publications and applications at about a 30 percent

[11] Renewal applications come from previously awarded projects that have reached funding expiration.

[12] The premium analyses include RFAs that do not explicitly state the amount of funds expected to be allotted because only realized award magnitudes are relevant for the analyses.

[13] Online Appendix A outlines these motivations in further detail based on interviews with NIH staff.

[14] Using the count of abstracts associated with specific MeSH terms follows the same logic as prior work that proxies for the scientific importance of particular genes with the number of publications related to each gene (e.g., Williams 2013).

larger rate than the control terms. RFAs tend to focus on topics more popular than average.[15] Still, the coverage of RFA-treated subjects is substantial.

To examine the potential for policy endogeneity, I assume that if RFAs are indeed endogenously created in response to prior events, then I should observe scientists pursuing RFA-treated MeSH terms at an increasing rate prior to the RFA; there will be "pretrends."[16] To empirically test for these trends, I use event study regressions to estimate the number of abstracts N associated with MeSH term m at time t in a conditional Poisson model:

$$(1) \qquad E\left[N_{mt} \,|\, RFA_{m\tau}; \beta_\tau, \delta_m, \gamma_t\right] = \exp\left(\sum_{\tau=t-\underline{s}}^{t+\bar{s}} \beta_\tau \times \mathbf{1}\{RFA_{m\tau}\} + \delta_m + \gamma_t\right),$$

where E is the expectation operator; δ_m and γ_t are MeSH and time fixed effects, respectively; and $\mathbf{1}\{RFA_{mt}\}$ equals one when MeSH term m is associated with at least one RFA at time t.[17] As in standard event study regressions, the coefficients of interest, β_τ, describe the rate of the dependent variable for time periods spanning $\underline{s}$ periods prior to t and $\bar{s}$ periods after t. In all models, β_τ is estimated for four years prior to t and two years post t, which corresponds to $(\underline{s}, \bar{s}) = (4, 2)$ for the PubMed data given annual observations and $(\underline{s}, \bar{s}) = (12, 6)$ for the NIH data given three observations per year.[18] Thus, $\beta_{\tau=t}$ estimates the relative change in term occurrence in the period solicited by an RFA. Figure 2 plots the β_τ estimates for three samples: PubMed publications, all NIH applications, and successful applications. The results clearly show that, conditional on the fixed effects, there are no significant pretrends.

Focusing on the NIH data, there is a sharp increase in applications using targeted MeSH terms at the time of the RFA ($\beta_{RFA} = 10$ percent) and awards ($\beta_{RFA} = 20$ percent).[19] RFAs appear to induce and fund applications that would not otherwise have been funded. Of note is the lack of any persistent post-RFA treatment effect. This suggests the response observed is due to RFAs specifically and not any spillovers.[20] These event studies support the main identification assumption of this paper: scientists value RFAs in and of themselves.[21]

[15] What this means for generalizability will depend on whether the costs of incentivizing scientists to pursue topics varies across this distribution. On one hand, the costs of pursuing low-rate subjects may be higher if the low rate is indicative of the low net value associated with those topics. On the other hand, pursuing the high-rate subjects may be more costly, as they are already concentrated with scientists and the marginal contribution may be more difficult.

[16] It is useful to note that the impetus for any RFA typically begins at the beginning of each fiscal year for budget purposes, if not sooner. Thus, if RFAs are responding to endogenous events, those events likely occurred in the years prior. And while, in general, such pretrends are neither necessary nor sufficient for arguing exogeneity of an event, this setting is one where it is very likely that differential trends would appear if RFAs were in fact endogenous to other unobserved events.

[17] Estimating this Poisson formulation handles the count nature of the data appropriately and the fact that certain MeSH terms have greater variance in the outcome simply because, for example, they are broad terms that encompass larger ideas. For example, "Neoplasms" and "Large Granular Lymphocytic Leukemia" are MeSH terms that describe any cancer and a very specific type of cancer, respectively.

[18] These three observations correspond to the three annual application rounds the NIH holds.

[19] The lack of a treatment effect in the PubMed data is likely due to both the much larger scale of publications compared to applications (average MeSH rate of 322 compared to 4) and the slow variable process by which NIH awards eventually give rise to new publications. Section IV's analyses focus more specifically on this question of whether RFA awards do in fact lead to new science.

[20] For instance, if RFAs were a strong signal of future funding opportunities at the NIH, then we would have expected to observe significant posttrends.

[21] While empirically useful, the lack of an apparent correlation between the supply/demand of the topics that are targeted by RFAs has interesting implications for the overall efficiency of these mechanisms. Discussing these

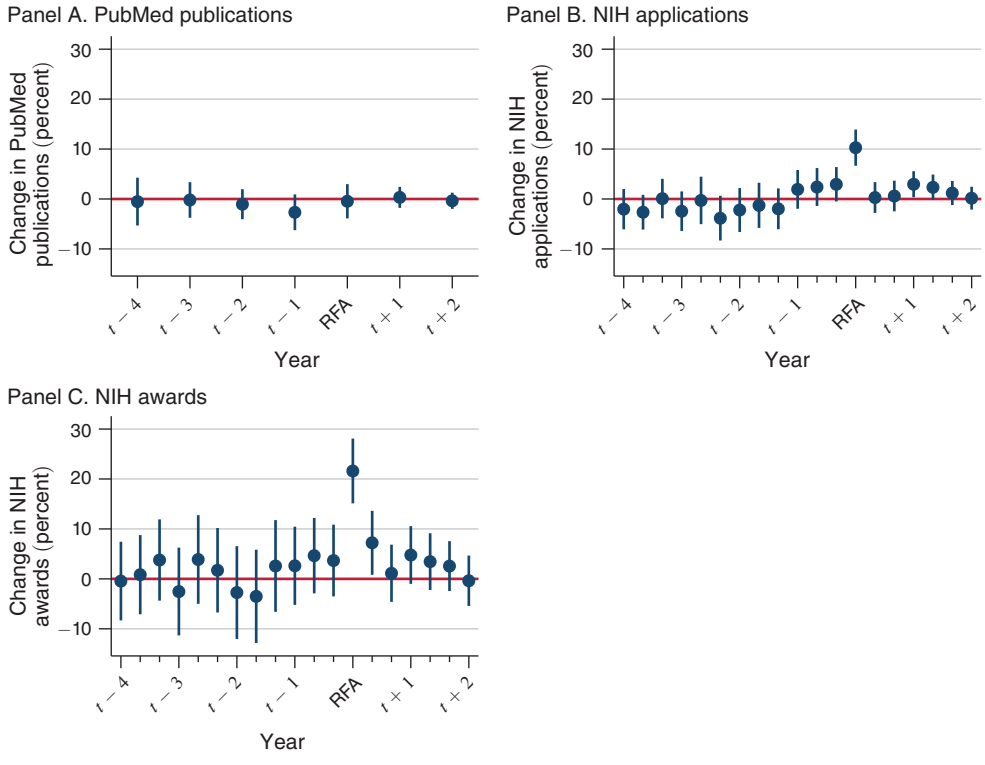

FIGURE 2. EVENT STUDY OF RFA-TARGETED SUBJECTS

Notes: This figure plots the β_τ coefficients estimated using equation (1). Ninety-five percent confidence intervals (based on standard errors clustered at the MeSH term level) are plotted as bars.

D. *Quantifying Direction with Similarity: The pmra Algorithm*

Fundamental to the notion of redirecting scientists is "how much" their course of work is adjusted. The task of moving a scientist from working on topic *A* to topic *B* will depend largely on the *scientific similarity* between *A* and *B*. Do they make use of the same knowledge? Do experiments use the same inputs, such as chemicals or organisms?

Here, scientific similarity is defined as the overlap in scientific terminology between two sets of text. Depending on the analyses, I compare two of the three following sources: publication abstracts, NIH application abstracts, and the "research objectives" section of RFAs. To estimate this similarity, I employ the most widely used similarity estimation algorithm for the biomedical sciences: the PubMed related articles (*pmra*) algorithm, developed by Lin and Wilbur (2007).[22] Online Appendix C outlines the algorithm in more detail and provides examples for

and other welfare issues surrounding absolute magnitudes (e.g., "which types of science should be targeted by RFAs?" or "how large should the total NIH budget be?") is beyond the scope of this paper.

[22] *pmra* is the algorithm underlying the "Similar articles" feature of PubMed and has become a benchmark within the field of bioinformatics for measuring similarity.

qualitative interpretation. The intuition behind the algorithm is that if two abstracts both use the same scientific terminology, it is likely the underlying science is more similar, especially if the overlapping terms rarely occur in general.[23]

The *pmra* algorithm requires PubMed publication abstracts as inputs, so I must assume that each scientist's knowledge base is embodied within the articles they have published previously. Indeed, the purpose of publication is to disclose generated knowledge. Thus, the number of similarity scores for each scientist-RFA or scientist-application pair equals the number of prior publications. To simplify this vector to a single value per scientist, I use the maximum.[24] The logic is that if publications define the boundary of each scientist's knowledge, the maximum captures the shortest proximity between the new science and the scientist. Given the likelihood of a general drift in scientists' pursuits, I use each scientist's publications within the most recent five years.[25]

Although qualitative interpretations of these similarity score changes are difficult, online Appendix C uses two examples from a set of biomedical and economics publications to make clear that changes in similarity scores observed in this sample are not dramatic in a qualitative sense. For example, one standard deviation in these data does not describe an biologist deciding between a new career as an economist or a piano player but rather a virologist deciding which of two closely related viruses to study for their next project.[26]

II. Entry Decisions and Elasticity Estimates

A. *Empirical Model*

The goal of this section is to understand the extent to which larger NIH investments can lead to more project proposals. And more generally, I am interested in quantifying the costs facing scientists looking to adjust their research direction. Understanding these field-to-field switching costs is important for devising optimal research policies.[27] The analyses use scientists' observed entry decisions to infer

[23] This algorithm has been used in similar work by Azoulay, Fons-Rosen, and Graff Zivin (2019). A novel feature of my implementation of the *pmra* software is that I can generate similarity scores between published journal articles and user-defined text. The code for my implementation of *pmra* was very kindly developed by W. John Wilbur of Lin and Wilbur (2007).

[24] For any applications listed with co-principal investigators (i.e., two or more scientists are project leads), I use the maximum of the two scientists' scores.

[25] Certainly, the density of one's knowledge may vary within these boundaries, but using the maximum ensures the measure captures only variation in similarity information rather than its depth (or quality). Robustness checks reported in online Appendix F using other transformations (i.e., mean, median), including the full set of prior publications or controlling for the number of publications, show no qualitative differences in the results.

[26] Purely for illustration purposes, the example in online Appendix C using economics publications shows that a 50 percent change in similarity scores amounts to the study of marginally different aspects of the pharmaceutical R&D industry (i.e., studying how firms use alliances in their R&D strategies versus studying how Medicare Part D impacted firms' R&D investments).

[27] For instance, in directed technical change models (e.g., Acemoglu 2002), a key parameter that determines the value of policy intervention (conditional on market size and prices) is "state dependence." This describes the extent to which prior factor-specific investments differentially influence current productivity. This parameter is partly based on the presence of across-factor switching costs. For instance, the optimal energy policies developed by Aghion et al. (2016) depend on how costly is it for firms working on "dirty" energy technologies to switch into working on "clean" energy technologies given the technological difference between those two pursuits. The

their preferences for more (or less) grant funds and less (or more) of a change in the nature of their work.

Motivated by the pattern in Figure 1, panel A, I estimate the probability that scientist *i* enters RFA *j* per

$$\mathbf{1}\{Entry_{ij}\} = F(Similarity_{ij}) + G(Purse_j) + \boldsymbol{\gamma}\mathbf{X}_j + \alpha_i + \epsilon_{ij}, \tag{2}$$

where $\mathbf{1}\{Entry_{ij}\}$ equals one if scientist *i* submitted an application to RFA *j* and zero otherwise, $Similarity_{ij}$ is the *pmra* score between *i* and *j*, $Purse_j$ is the amount of funds made available in the RFA *j*, $\mathbf{X}$ is a vector of other RFA-specific characteristics,[28] α are scientist fixed effects, and ϵ is a mean zero error term.

Under the assumption that the *Purse* and *Similarity* variables are orthogonal to the error term (conditional on $\mathbf{X}$ and α), estimating this model identifies the marginal costs (to the NIH) of attracting additional applicants to an RFA. And the findings of Section IC support this assumption. These cost estimates are relevant for NIH officials debating the costs and benefits of intervening in the allocation of public research funds.

Furthermore, some portion of these costs will be driven by the elasticity of science. In fact, if equation (2) is interpreted as a structural model of scientists' payoffs, then the ratio of G' to F' (i.e., the coefficients on the *Purse* and *Similarity* variables) can be used to solve for this elasticity (ignoring subscripts and abbreviating *Similarity* and *Purse* to *S* and *P*, respectively):[29]

$$\frac{G'}{F'} \cdot \frac{P}{S} = \frac{\partial \Pr(\mathbf{1}\{Entry\} = 1)/\partial P}{\partial \Pr(\mathbf{1}\{Entry\} = 1)/\partial S} \cdot \frac{P}{S} = \frac{\partial S}{\partial P} \cdot \frac{P}{S}, \tag{3}$$

where the final elasticity term indicates how much a scientist can be incentivized to change their direction given some expectation of funds. However, the G' recovered from estimating equation (2) as is would not directly reflect scientists' preferences for grant funds per se. The issue is competition. Larger purses—the sort of variation used to identify G'—will likely lead scientists to have higher expectations about the number of competitors they will face. If this is the case, then the relative weight scientists will appear to place on these funds will be biased downward to zero.[30]

To identify these more general switching costs of science, I must control for competitive expectations. To do so, I implement a method for estimating static strategic interactions outlined by Bajari et al. (2010), which builds on the "two-step" method of estimating games pioneered by Aguirregabiria and Mira (2007). This approach amounts to first estimating the probability that each scientist enters each RFA using the exogenous covariates. Then these entry probabilities—often referred

elasticity I estimate captures the same force—how costly is it for scientists to change the direction of their work given the magnitude of that change.

[28] These include the year the RFA expired (typically six to eight months post announcement), the lead NIH Institute responsible for the RFA, whether non-R01 applications were permitted, and whether the RFA specifically requested "collaborative" research be proposed. All of these covariates enter as fixed effects (i.e., year and NIH Institute dummies, etc.).

[29] See online Appendix D for a more specific discussion.

[30] The magnitude of this bias depends on the specific nature of competition in these RFAs.

to as conditional choice probabilities—are used to construct a measure of competitive expectations, which is included as an additional covariate in the entry model.

For simplicity and tractability given the large sample, though at some costs discussed below, I perform these two steps using the following linear regression models:

$$\text{(4a)} \quad \mathbf{1}\big(Entry_{ij}\big) = \tilde{F}\big(Similarity_{ij}\big) + \tilde{G}\big(Purse_j\big) + \tilde{\boldsymbol{\gamma}}\mathbf{X}_j + \tilde{\alpha}_i + \tilde{\epsilon}_{ij},$$

$$\text{(4b)} \quad \mathbf{1}\big(Entry_{ij}\big) = F\big(Similarity_{ij}\big) + G\big(Purse_j\big) + \delta\tilde{n}_{ij} + \boldsymbol{\gamma}\mathbf{X}_j + \alpha_i + \epsilon_{ij}.$$

Estimates of the parameters in equation (4a) $(\tilde{F}, \tilde{G}, \tilde{\boldsymbol{\gamma}}, \tilde{\alpha}_i)$ are used to predict the entry probabilities, which are in turn used to construct the competitive expectations variable denoted by $\tilde{n}_{ij}$.[31] This variable describes the number of competitors scientist *i* expects to face in RFA *j*. Online Appendix D outlines the procedure in further detail.

Identification of this class of models rests on the so-called "strategic exclusion restriction." The instrumental variables that influence each scientists' entry decisions directly—here, similarity and the scientist-specific intercepts—must only influence other scientists' decisions indirectly via their competitive expectations $(\tilde{n}_{ij})$. If this assumption holds, then conditioning on this covariate (as in equation (4b)) serves the very useful purpose of holding competition fixed. With competition fixed, variation in entry decisions driven by purse sizes reflects scientists' preferences for grant funds per se, and the elasticity of science can be estimated.

The implications that stem from this assumption and the structure of equations (4a) and (4b) are not trivial. Namely, I assume that all scientists are fully aware of all other scientists and their similarities to each RFA. And I do not allow for any heterogeneity in the responses to different types of competition; all scientists respond to competition from all other scientists in the same way (as captured by the linear term δ in equation (4b)). While these are strong assumptions, they allow me to adjust for competitive expectations in a clear way. Since the goal of this analysis is not to estimate any sort of counterfactual with a fully specified model, the costs of these strong assumptions seem worth the extra insight they provide.[32]

I explore alternative functional forms for F and G. In many specifications, I replace $F(Similarity_{ij})$ with $\sum_b \beta_b \big(\mathbf{1}\big\{bin\big(Similarity_{ij}\big) = b\big\}\big)$, where $\mathbf{1}\{\cdot\}$ is the indicator function and the bins b are discretized groups of the *pmra* similarity score. And to explore heterogeneity with respect to the effect of purse sizes on different scientists, I also estimate a version of the model where purse size is interacted with

[31] $\tilde{n}_{ij} = \sum_{i' \neq i}^{N} \tilde{\Pr}\big(\mathbf{1}\big\{Entry_{i'j}\big\} = 1\big)$, where N is the total number of potential entrants, and the term encompassed by $\tilde{\Pr}$ is the predicted entry probabilities based on the estimates from equation (4a).

[32] This research design resembles studies of firm entry where certain variables, such as geographic distance, generate exogenous variation in entry (e.g., Krasnokutskaya and Seim 2011). Here, although the *pmra* algorithm provides an estimate of scientific distance, it is unclear whether scientists must actually travel the entirety of the distance since the bounds of an RFA are not explicit. In online Appendix C, I outline a process that adjusts the raw RFA-scientist similarity estimates to account for this fact. Robustness tests of the main results show no qualitative changes without this adjustment. The adjusted *pmra* is preferred because it more accurately captures the real changes to be expected by entrants, and it permits a straightforward comparison to the effects identified in Section III, where realized changes in similarity are examined.

these similarity score bins. This informs how the costs of attracting scientists may depend on how "close" they are to a particular RFA.

B. *Data Construction and Summary Statistics*

Estimating this model requires a dataset composed of scientist-RFA pairs containing (i) all potential entrants, (ii) RFA-level data on purse size and research objectives, and (iii) each potential entrant's scientific similarity to the research objectives. To arrive at a close approximation to the full set of potential entrants, I first include any individual that applied to the NIH from 2002 to 2015, totaling 142,745 scientists. I then pair each scientist (and their publication history) with each of the RFA announcements between 2002 and 2009 that solicit R01 research grants and for which details on the timing, administration, research objectives, and purse were available. Dropping scientists that receive *pmra* scores of zero for all RFAs results in a $118{,}127 \times 386$ observation dataset of RFA-scientist pairs and is the data underlying Figure 1.

But as illustrated in Figure 1, panel A, there is virtually no meaningful change in entry probabilities below the median similarity value. From the median to the seventy-fifth percentiles, there is minimal change, but afterward there begins a very sharp increase. Based on this pattern, I first trim the data to exclude all scientists from below the median similarity score. This eliminates roughly 15 percent of applicants (roughly 1,800 out of 13,000). However, the focus of this analysis is understanding how the majority of scientists make directional decisions. The fact that these particular applicants appear to be extreme outliers, coming from regions far outside where the average applicant comes from, suggests that the loss in information from excluding them from the analysis is worth the gains in precision when focusing on the upper half of the full distribution.

The construction of the similarity bins b used to flexibly approximate F is also informed by Figure 1, panel A. I set the reference group of scientist-RFA pairs to be those within the fiftieth to seventy-fifth percentiles of the full distribution and then create 25 one-percentile bins for the remainder of the observations. Thus, the β_b coefficients will capture the relative change in entry probability for scientist-RFA pairs in the bth percentile ($b \in [76, 100]$) of similarity scores relative to those in the fiftieth to seventy-fifth percentile.

In this final sample, the average RFA has a first-year total purse size of \$2.89 million (standard deviation (SD) = 2.65), attracts 28.7 entrants (SD = 24.4), and awards a total of 6.5 grants (SD = 5.5), with first-year direct costs totaling \$2.65 million (SD = 3.01).

I lack an identification strategy that allows me to conduct a detailed analysis of differences in how scientists compete conditional on entry.[33] However, to get a sense as to the competitiveness of scientists whose prior work is more or less aligned with the RFA objectives, Figure 3 plots the similarity distributions for winners and losers. Winners are more likely to be more similar to the RFA, but the shift in the

[33] This would require a second instrument, as the *pmra* scores are likely predictive of entry both based on the fixed costs of entry and the expected benefits conditional on entry.

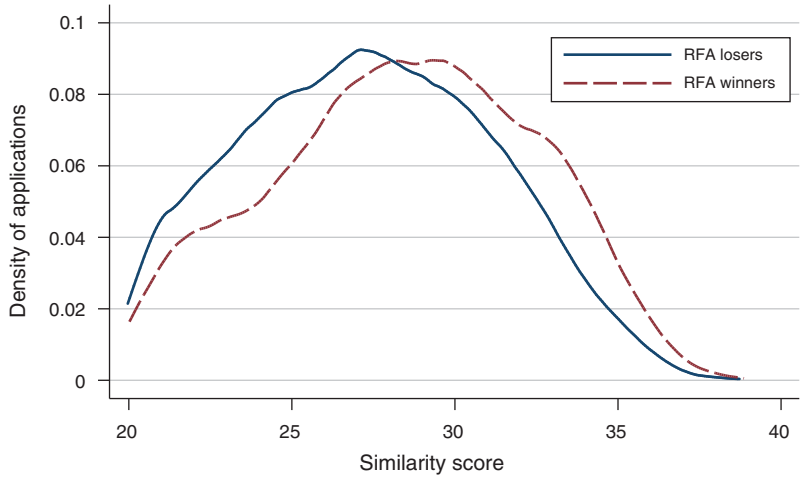

FIGURE 3. SIMILARITY DISTRIBUTIONS CONDITIONAL ON ENTRY

Note: This figure plots the adjusted *pmra* similarity scores between each scientist and the RFA applied to, splitting the sample by winners (received a grant) and losers.

distribution is not substantial. Unreported regressions suggest this difference is on the order of only 2 to 5 percent. Conditional on entry (which likely includes expectations about success), there does not appear to be a dramatic difference in scientists' ability to compete.

C. *Results*

Table 1 presents the main results from estimating variants of equation (2). The reported coefficients are based on standardized transformations of the independent variables, so all can be interpreted as the change in entry probability associated with a one standard deviation increase in the variable. The first three columns control for the scientist-RFA similarity using the flexible bin structure and focus on the role of the purse size. Across the specifications, where varying degrees of RFA- and scientist-level controls are included, the estimates indicate that a one standard deviation ($3 million) increase in funds allocated to an RFA would increase the probability that each scientist enters that RFA by about 40 percent relative to the mean.

These coefficients imply that the marginal costs, in terms of purse size, of attracting one additional application from this sample are roughly $110,000–120,000. Given that purse sizes are measured only as a single year of funding that typically lasts four years, this implies net costs of roughly $450,000, which is nearly 40 percent of the lifetime size of the average R01 grant.

In a perfectly competitive market of RFAs with no frictions, adjustment costs, or any other constraint, it would be expected that the marginal cost of inducing an additional application would be equivalent to the expected value of entry. The expected value of entry is simply the probability of winning times the expected award size. Unfortunately, I cannot identify win probabilities for marginal applications within

Table 1—Determinants of RFA Entry

	$\mathbf{1}\{Entry_{ij}\}$				
	(1)	(2)	(3)	(4)	(5)
$Purse_j$	2.14	2.20	2.25	2.32	4.07
	(0.557)	(0.515)	(0.519)	(0.551)	(0.503)
$Similarity_{ij}$				2.33	2.55
				(0.911)	(0.964)
$Competitive\ Expectations_{ij}$					−4.37
					(0.271)
Includes similarity bins	Y	Y	Y		
RFA controls		Y	Y	Y	Y
Scientist fixed effects			Y	Y	Y

Notes: All models include 20,221,541 scientist-RFA (ij) pair observations, where the mean entry probability is 5.47×10^{-4}. Independent variables are standardized in regression, so coefficients indicate the change in entry probability associated with a one standard deviation increase in the variable; all coefficients are scaled by 10^{-4}.

RFAs. I do not have an instrument that influences entry but is orthogonal to win probabilities. Still, the average win probability of entrants in the sample is 21 percent. If I perform an RFA-level regression of the total number of awards on the total number of applications, marginal applicants across RFAs win 14 percent of the time.[34] Using 14 percent as the most reasonable and conservative estimate I can infer from the data, it would appear that the expected value of entry for marginal entrants is closer to $150,000, about 3 times less than the implied costs of soliciting additional applicants ($450,000). This suggests there are sizable costs keeping scientists from changing their direction and entering RFAs.

To get an initial sense of the relative importance of funding versus the scientific similarity of a scientist's prior work, column 4 includes a simple linear term and indicates that a one standard deviation increase in scientist-RFA similarity increases entry probabilities by more than 40 percent. Next, I take a more flexible approach to specifying the role of similarity and purse sizes (F and G). I use the similarity percentile bins described previously to estimate the relative effects of similarity, and I also interact these bins with the purse variable to explore how fund responsiveness varies across the similarity distribution:

$$\begin{aligned}(5)\qquad \mathbf{1}\{Entry_{ij}\} &= \sum_b \beta_b\Big(\mathbf{1}\{bin(Similarity_{ij}) = b\}\Big) \\ &\quad + \sum_b \delta_b\Big(\mathbf{1}\{bin(Similarity_{ij}) = b\}\Big) \times Purse_j \\ &\quad + \delta\tilde{n}_{ij} + \boldsymbol{\gamma}\mathbf{X}_j + \alpha_i + \epsilon_{ij},\end{aligned}$$

[34] In these scenarios, there is often concern that marginal entrants are "worse" than average entrants. But one feature to reiterate about this setting is that the size of RFAs is very small relative to the full pool of scientists—there are roughly 28 applicants per RFA out of more than 100,000 scientists. So it seems reasonable to assume that the difference between average and marginal entrants here is likely not dramatic.

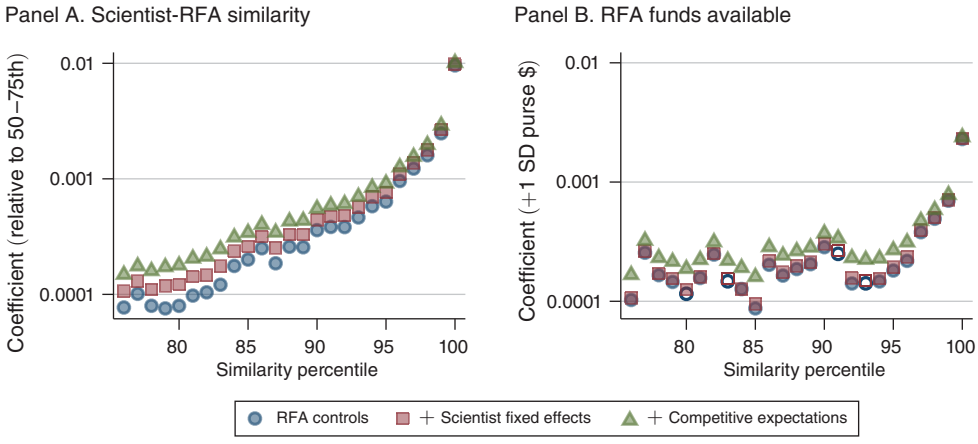

FIGURE 4. RFA ENTRY COEFFICIENT ESTIMATES

Notes: Panel A plots the coefficients on the similarity percentile bins representing the relative change in entry probability for scientist-RFA pairs in each percentile relative to those in the fiftieth to seventy-fifth percentiles. Panel B plots the coefficients on the standardized purse size interacted with each of the seventy-sixth to one hundredth percentile similarity bin indicators. All estimates are from one of three specifications indicated by the markers. Hollow markers indicate estimates where zero is included in 95 percent confidence intervals.

where the scientist fixed effects (α_i) and competitive expectations term $(\tilde{n}_{ij})$ are included only in some models. Figure 4 plots the β_b and δ_b coefficients recovered from this regression. The estimates in Figure 4, panel A indicate relative changes in entry probabilities for scientists in the *b*th similarity percentile relative to those in the fiftieth to seventy-fifth percentile. The estimates in Figure 4, panel B indicate absolute changes in entry probabilities when purse size increases by one standard deviation for scientists in the *b*th similarity percentile.

Focusing on Figure 4, panel A, the importance of similarity in scientists entry decisions is clear. The effect of being in the top percentile of similarity (approximately 10,000 individuals) is roughly 100 times larger than being at the eightieth percentile. This pattern holds across all specifications.

These results suggest that relative changes in similarity are much more important than fund availability. But as outlined previously, taking this pattern to indicate the presence of high switching costs ignores the potential role of strategic interactions. Column 5 implements the Bajari et al. (2010) algorithm to construct an estimate of competitive expectations (the number of entrants each scientists expects) and includes this term in the regression. Controlling for these strategic interactions leads to a larger coefficient on the purse term; it nearly doubles. This is to be expected since increases in fund availability should endogenously lead to increased competition, which mutes the responsiveness of scientists to these funds.

Figure 4, panel B illustrates how scientists' responsiveness to funds varies across the similarity distribution. For scientists outside the top 5 percent of similarity, there appears to be a relatively constant effect of purse size on entry decisions—consistent with Table 1, a 1 standard deviation increase in fund availability leads to a 40–80 percent increase in entry probability. However, it appears that scientists in the top

5 percent of similarity are significantly more responsive to funds. For these individuals, that same change in funding increases entry probabilities by two- to threefold. These patterns indicate that the scientists most responsive to RFAs are likely to be already operating in the area of science targeted. Still, distant individuals do exhibit some degree of responsiveness.

A number of robustness tests are reported in online Appendix F. I estimate equation (2) using unadjusted *pmra* scores; the max, median, and average of *pmra* scores based on the full set of scientists prior publications; and log transformations of the key independent variables. All results are very similar to the main results reported here.[35]

I can make more precise statements about the elasticity of scientists' direction if I interpret the similarity scores as cardinal measures of scientific space that scientists must traverse. Figure 5, panels A and B recreates Figure 4, panels A and B but converts the coefficient estimates into elasticities that indicate the percentage change in entry probability caused by a percentage change in similarity and purse size, respectively. For all but the most distant scientists, relative changes in similarity scores have much larger effects than relative changes in purse sizes, even when competition is held fixed. But both exhibit a pattern very similar to the RFA where responsiveness is particularly larger for scientists.

As outlined above, the ratio of these two elasticities describes the increase in funding necessary to induce a relative change in direction—the elasticity of science. Focusing on the specification that conditions on competitive expectations, Figure 5, panel C plots the ratio of these elasticities. Across the similarity distribution, I find the ratio of these two effects to be between around 0.1, which suggests that a 10 percent change in funding is necessary to induce a scientist to undertake a 1 percent change in their direction. Despite the nonlinearities documented in Figure 5 panels A and B, this elasticity is relatively consistent across the similarity distribution. I take this as evidence that the ratio of these two forces—similarity and funds—does in fact capture the underlying presence of directional adjustment costs facing scientists.

To convert this estimate into a more policy-relevant number, I do the following. First, I scale purse sizes by four since they describe only the first year's worth of funds and the average R01 lasts roughly four years. Second, I use the conservative estimate of win probabilities of 14 percent to convert purse sizes into expected values. Using these two adjustments, Figure 5, panel D plots the amount of grant funds needed in expectation to induce a scientist to undertake varying degrees of adjustment to their direction. I use the mean elasticity of science (approximately 0.1) to make the main estimates (the center line) and bound these estimates (with the shaded area) using the minimum and maximum elasticities.

For a reference point, a one standard deviation redirection in this sample corresponds to a 27 percent change in similarity. That is to say, relative to the mean similarity between scientists' prior work and an RFA, scientists who are one standard deviation closer to the RFA had 27 percent larger similarity scores. The results imply that incentivizing a scientist to undertake this level of redirection would require a

[35] Unreported results using a fourth-order polynomial expansions of the competitive expectations term are also very similar to the main results.

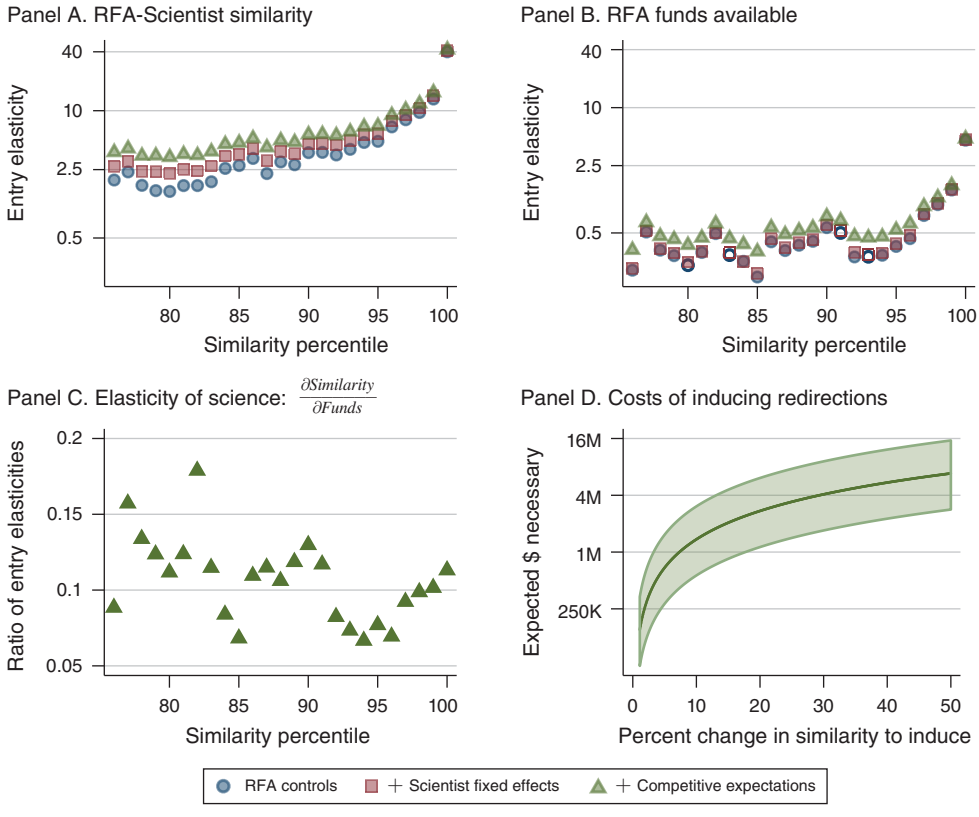

Figure 5. Implied Elasticities and Inducement Costs

Notes: Panels A and B plot the implied percentage change in entry probability given a 1 percent change in similarity or fund availability at the sample means; note the log scale of the *y*-axes. Panel C plots the ratio of these elasticities to give the implied percentage change in similarity that can be induced by a 1 percent change in funding. All estimates are from one of three specifications indicated by the markers. Hollow markers indicate estimates where zero is included in 95 percent confidence intervals. Panel D projects the implied redirection costs as described in the text.

total of about $3.7 million in expectation, nearly three R01s worth of funds. In the full sample of *pmra* scores, one standard deviation corresponds to a 50 percent change in similarity. Extrapolating my estimates suggests that inducing this degree of change would require approximately $6.8 million in expectation.

Certainly, it is easy to find examples of scientists making large changes in their direction. The problem with inferring the costs of these movements is that scientists follow the most promising new opportunities. The estimates from these analyses, which hold fixed these opportunities, indicate that the costs of incentivizing marginal directional changes are very large relative to the amount of funds scientists can ever expect from a single NIH grant.

III. Is There an RFA Premium?

The goal of this section is to investigate whether the costs of redirections identified in the prior section affect the equilibrium size of grants. If it is costly for

FIGURE 6. TRENDS IN OUTCOMES—RFA VERSUS OPEN APPLICATIONS

Notes: RFA-open correlations are 0.97 for similarity, 0.90 for win probability, and 0.92 for award size. The lack of abstract text for pre-2007 restricts the set of applications for which similarity scores can be calculated.

scientists to adjust the direction of their work, and RFAs require larger adjustments relative to the open grants, then the expected value of RFAs could be larger. Figure 6 illustrates that, compared to open applications, RFA applications are 25 percent less similar to the scientist's prior work, on average.[36] Just as compensating differentials arise when occupations have undesirable attributes, so too may RFAs be larger than open grants as they require scientists to undertake costly redirections.[37]

The main conditions underlying this argument are that competition reduces the expected value of applying and marginal applicants are indifferent between the costs and benefits of applying to RFAs or open grants—the market is in equilibrium. To the point of competition, NIH policies make clear that the division of funds takes

[36] Online Appendix C shows the full distribution of RFA similarity scores, which is shifted to the left (less similar) compared to open applications.

[37] Online Appendix E provides a simple, formal treatment of this argument, which is a discrete version of the experiment from the previous section. Instead of choosing from RFAs that each require a different degree of redirection, scientists now only choose between two options: RFAs or open grants.

into account the number and quality of applications.[38] More competition should reduce expectations about award size. And as evidenced in Figure 6, the expected payoff of RFA and open grants trend together very closely over time. This suggests that, each year, applicants evaluate the costs and benefits of pursuing particular grants and choose the option with the largest payoff.

The goal of the following regressions will be to identify whether these differences are causally due to RFAs or if they are an artifact of selection bias (e.g., scientists with high-quality ideas are more willing to redirect, are more likely to win, and win larger awards). Furthermore, if RFAs cause marginal scientists to change the direction of their work and in turn have a larger expected value, I can use the elasticity estimates from the previous section to explore how much of this RFA premium can plausibly be explained by scientists' (un)willingness to switch topics.

A. *Empirical Model*

As a first pass, I could regress the three outcomes $y = \{$scientific similarity, win probability, award size$\}$ for application j submitted by scientist i on an indicator for whether the application was submitted to an RFA ($\mathbf{1}\{RFA_j\}$) as follows:

$$y_{ij} = \alpha + \beta \mathbf{1}\{RFA_j\} + \epsilon_{ij}. \tag{6}$$

From these regressions, the β coefficient can be used to describe the RFA redirection effect (change in similarity) and premium (change in expected value).[39] But this simple difference suffers from two drawbacks: (i) it is unclear that the average open application is the appropriate counterfactual for RFA applications, as is implied by equation (6); and (ii) scientists likely have expectations about their potential outcomes in either competition, which could generate a selection bias.

To the first point, I introduce scientist fixed effects to remove any variation driven by stable differences across individuals. I also use the funding channels and peer review groupings described in online Appendix A to group all applications (RFA and open) into highly detailed sets of research areas. This process generates roughly 400 different research areas, denoted s, each with an average of 40 applications per year. Furthermore, I interact these research areas with time fixed effects to generate a rigorous set of research area–time fixed effects. Including these fixed effects in the most saturated (and therefore conservative) model assumes that, conditional on submitting an RFA (or open) application, the most similar outcome for that individual would have been to compete within the open (or RFA) competition of the same research area that same year.

[38] Case in point is this language that appears in virtually all RFAs in some way or another: "The total amount awarded and the number of awards will depend upon the mechanism numbers, quality, duration, and costs of the applications received." This particular quote is from the following RFA: https://goo.gl/wZPmij (accessed July 12, 2017).

[39] The redirection effect will be directly estimated as $\beta_{Similarity}$. I combine the β coefficients from the win probability and award-size regressions in a rational expectation framework where the expected value of entry is the product of the two. Letting $\bar{w}$ and $\bar{a}$ denote the average win rate and award size in open grants, respectively, the shift in expected value is then simply of the form $(\bar{w} + \beta_{Win}) \times (\bar{a} + \beta_{Award})/(\bar{w} \times \bar{a})$.

To control for the possibility of selection bias beyond the fixed effects, I condition on the peer application's review score and the amount of funds requested, which plays a large part in determining the size of the award. Controlling for these variables accounts for scenarios where, for example, scientists may be more willing to undertake costly redirections into RFAs if they believe they have a high-quality proposal that is more likely to be funded. Conditioning outcomes on the amount of funds requested is also useful because it ensures that applications of equivalent scope/scale are compared to one another.[40]

Rewriting equation (6) to include the fixed effects for individuals (i) and time-variant research areas (st) and the covariates ($\mathbf{X}$), the main estimating equation is

$$y_{ijst} = \alpha_i + \beta \mathbf{1}\{RFA_j\} + \boldsymbol{\gamma}\mathbf{X}_{ij} + \tau_{st} + \epsilon_{ijst}. \tag{7}$$

Equation (7) is estimated using ordinary least squares (OLS) for the three outcomes: log-transformed similarity scores, a binary win indicator, and log-transformed first-year direct costs.[41]

B. *Summary Statistics and Results*

For the focal sample of applications included in this analysis, the average application is successful 23 percent of time, with an average first-year award size of $298,000. At its largest, the sample includes 26,734 unique scientists and 39,756 applications, with each scientists submitting an average of 1.5 applications (2.5 for those with greater than one application).

To investigate how much RFAs induce scientists to adjust the trajectory of their research relative to open grants, Table 2, panel A reports the results of estimating equation (7) with the *pmra* similarity score as the dependent variable. In all specifications, there is a persistent difference in the degree to which each scientist's new application resembles their prior work; RFA applications are less similar. The magnitude of this difference declines upon the inclusion of the controls, mostly driven by the individual and research area fixed effects, suggesting that there are important underlying differences across scientists and fields of research with respect to how individuals' research trajectories evolve. Overall, the evidence indicates that RFAs do force scientists to pursue a topic for funding that is less similar on average than what they would have otherwise pursued with an open grant. The models with scientist fixed effects suggest this difference is roughly 10 to 15 percent.

Table 2, panels B and C present estimates of the win probability and award-size differences. Across all specifications, RFA applications are more likely to be awarded funds and, conditional on winning, are likely to receive a larger amount of funds. Combining these two effects, the magnitudes indicate an RFA premium of

[40] Including this as a control means that estimated differences in award size are, with respect to the "surplus" of award funds, relative to the amount requested.

[41] Online Appendix F includes results using alternative independent variable transformations and versions of award size (i.e., total costs).

Table 2—RFA versus Open Applications

	(1)	(2)	(3)	(4)	(5)
Panel A. $\log(Similarity_{ijst})$					
$\mathbf{1}\{RFA_j\}$	–0.236	–0.242	–0.153	–0.110	–0.101
	(0.0089)	(0.0089)	(0.0093)	(0.0157)	(0.0185)
$N_{Applications}$	39,741	39,731	21,495	36,237	18,582
$N_{Scientists}$	26,723	26,716	8,480	24,864	7,408
mean$(Similarity_{open})$	33.5	33.5	34.3	33.4	34.3
Panel B. $\mathbf{1}\{Win_{ijst}\}$					
$\mathbf{1}\{RFA_j\}$	0.111	0.0749	0.0706	0.121	0.0883
	(0.0065)	(0.0050)	(0.0087)	(0.0096)	(0.0182)
$N_{Applications}$	39,756	39,746	21,502	36,250	18,588
$N_{Scientists}$	26,734	26,727	8,483	24,874	7,411
mean$(\mathbf{1}\{Win\}_{open})$	0.22	0.22	0.21	0.22	0.21
Panel C. $\log(Award\$_{ijst})$					
$\mathbf{1}\{RFA_j\}$	0.240	0.199	0.197	0.0587	0.167
	(0.0149)	(0.0112)	(0.0229)	(0.0235)	(0.0504)
$N_{Applications}$	9,743	9,738	1,944	8,447	1,182
$N_{Scientists}$	8,698	8,693	899	7,605	555
mean$(Award\$_{open})$	270,646	270,670	288,053	271,693	284,237
Application controls		Y	Y	Y	Y
Scientist fixed effects			Y		Y
Research area–time fixed effects				Y	Y

Notes: Application controls include peer review score and funds requested. Standard errors are clustered within scientists. *Award*$ refers to first-year direct costs of the grant.

roughly 65 percent relative to the expected payoff of open applications. This implies that scientists are indifferent between approximately $270,000 (in first-year funds) awarded via an open grant and $445,000 awarded via an RFA.

At first glance, this seems like a large premium to place on RFAs. But given the magnitude of redirection they appear to require, this plausibly could be explained by the (in)elasticity of science identified in the prior section. To explore this possibility, I must assume that the entirety of the RFA premium is due to this redirection effect. If that were the case, then these results imply that scientists are indifferent between a 10 percent redirection to more similar work and a 65 percent increase in expected grant funds. These magnitudes imply an elasticity of direction with respect to grant funds of about 0.15 $(= 0.1/0.65)$. For comparison, the average elasticity identified with the entry model was approximately 0.11.

These two estimates are quite similar, suggesting that a significant portion of the RFA premium could be explained by the switching costs facing scientists considering these mechanisms. There are two caveats to this suggestive finding. The first is that I cannot perfectly attribute the RFA premium described by the results in Table 2, panels B and C to the redirection effect in panel A and not to any other RFA-specific factors that scientists may value positively or negatively.[42]

[42] As described earlier, the two funding channels differ in the nature of competition, with open competitions being much larger and broader in scope. Furthermore, the NIH allows scientists to attempt to extend the duration of a grant beyond its initial timeline, a process referred to as "competitive renewal," and anecdotal evidence indicates

Second, the elasticity identified by the entry model includes (essentially) all potential applicants, while the analysis in this section examine outcomes only for applicants. Are these the same marginal applicants? It is reasonable to assume that if individuals are on the margin for participating in an RFA, then they would also consider participation in an open mechanism, and vice versa. And the inclusion of the scientist fixed effects removes stable, unobservable variation across scientists. However, it is possible that scientists who do apply for an NIH grant are doing so at particular times in their careers when their demand for NIH funds is higher. Perhaps a prior grant has expired, perhaps their institution has implemented budget cuts, etc. To this point, it is not surprising, then, that I find applicants to be slightly more elastic than the larger pool of potential applicants on average. Still, I take this as new evidence that the adjustment costs facing scientists can have a first-order effect on the equilibrium size of grants awarded by the NIH.

IV. Grant Productivity

The results of Sections II and III are based on the information contained within scientists' grant applications, their *intentions*. But it is by no means guaranteed that, if awarded, grants ultimately influence the rate or direction of scientists' work.[43] Furthermore, the efficiency of RFAs will depend on whether the productivity of grants differs between the RFA and open mechanisms. This analysis builds on Azoulay et al. (2019) and Jacob and Lefgren (2011), which examine the science- and scientist-level impact of NIH funding. My focus is on comparing the scientist-level effect of open and RFA grants. RFAs may lead to fewer publications than open grants if, for example, winners who are inexperienced in the specific area targeted by the RFA are not able to deliver on what they propose. Conversely, if RFAs encourage work in "hot" areas, or if the fresh perspective of being a new entrant to the field increases productivity, RFAs may in fact lead to more publications per grant (or dollar).

The following empirical approach is closer to Jacob and Lefgren (2011), which focuses only on the open mechanism and uses a fuzzy regression discontinuity design. In short, they use the imperfect rank-order funding process whereby the NIH sorts applications per their peer reviewed quality and (probabilistically) funds applications from top to bottom per their budget constraints. Because this process is not perfectly due to rank order, the authors must generate pseudodiscontinuities to generate identifying variation in awards.[44] However, this imperfection in the rank-order award process is largely due to the potentially nonrandom preferences of the NIH administrators, complicating the interpretation of this variation. Instead, I exploit the fact that I can observe the budget requests of each application, which are both

that this option is much easier for open grants compared to RFA grants. I lack data on these renewal grants, so I cannot test for any differences.

[43] The key issues are that (i) NIH grant funds are relatively fungible across projects, so it is possible that the funds awarded in an RFA on topic *A* are in fact used for future research on topic *B*; and (ii) the market for grant funding of the top scientists may be competitive enough that losers at the NIH can still secure funding elsewhere.

[44] See Jacob and Lefgren (2011, 1171) for more. Note their footnote 10, which states: "Ideally, one would like to create the theoretical cutoff score taking into account the amount of funding associated with each application," which motivated the research design implemented here.

plausibly uncorrelated with NIH managers' preferences and, as will be shown, play a large role in determining whether *other applications* are awarded grants.

A. *Data and Dependent Variables*

Publication Rate.—I again use the Torvik and Smalheiser (2009) PubMed database to identify the set of publications each applicant is responsible for both before and after the award decisions, focusing on publications within five years of the award decision since most NIH grants last about four years. I identify publications where the scientist was a primary investigator by proxying for this role based on whether that scientist is listed as the first or last author on a publication. In the biomedical sciences, this is a very strong signal that an individual is the head of the laboratory or chiefly responsible for the design of the study.[45]

Publication Direction.—To assess changes in the direction of each scientist's publications—for the RFA grants only—I use the *pmra* algorithm to score the similarity between the applicant's publications (before and after award) and the research objectives of the RFA applied to. Again, larger scores indicate that the publication is more similar to the research objectives of the RFA.

B. *Empirical Model*, *Identification*, *and Covariate Selection*

A simplified overview of the award process is as follows: first, applications receive quality scores from peer review panels; second, applications are sorted based on these scores within their respective funding group;[46] finally, funds are awarded in imperfect rank order at the discretion of NIH staff until the budget constraint binds.[47] As just discussed, this discretion in the award process prevents clear discontinuities from existing in the data.

However, budgets are constrained, and so the goal of the following exercise is to separate the budget constraint effect from any NIH-preference effect.[48] I do so by leveraging the fact that I can observe the budget requests of each application and thus can calculate how "far away" from the best-ranked application each other application is. The logic is that, conditional on an application's peer review score, which mechanically determines its rank, the sum of budget requests for better-ranked applications is orthogonal to any feature of the focal application. However, this sum of budgets does influence the degree to which the budget constraint is binding. And

[45] To proxy for quality, I also explore specifications examining the average journal impact factor (JIF) of publications. The impact factor of a journal is the number of citations received in a given year of articles published in the journal in the two preceding years divided by the total number of articles published in that journal during the two preceding years. It is thus a noisy, but publicly available, measure of citations.

[46] For RFAs, this is simply within the RFA itself. For the open mechanism, each fiscal year applications are ranked within the institutes/centers they applied to for funding.

[47] Online Appendix A details the NIH award policies in further detail.

[48] Note that NIH preferences are not necessarily positively correlated with publication potential, since they do publicly commit to spreading funds to scientists across a wide range of geographical and institutional environments given the political context they operate in.

since this measure is based on funds requested but not necessarily awarded, it is not influenced by NIH staff in any way.

Consider two perfectly equivalent applications $j = \{1,2\}$ that happen to be in different funding groups. If the budget requests of the *other* better-ranked applications in $j = 2$'s funding group were much larger than those in $j = 1$'s group, then $j = 2$ would be less likely to receive an award. This is simply because managers evaluating application $j = 1$ would face a weaker budget constraint.

I calculate a scaled version of this budget-distance metric z for each application j in each RFA/open-funding group k, where k identifies each unique RFA or institute/center fiscal year grouping, as follows:

$$z_{jk} = \begin{cases} 0, & \text{if } rank_{jk} = 1; \\ \dfrac{\sum_{\bar{j}k} \$Request_{\bar{j}k}}{\$Request_k}, & \text{otherwise,} \end{cases} \tag{8}$$

where $\bar{j}$ denotes the applications in group k that are better (lower) ranked than application j, and $\$Request_k$ is the total amount of funds requested in the group k. This scaling is done to account for the fact that the number of applications, and thus total amounts requested, are endogenous to the budget constraint of each grouping. Being \$10 million away from the top application is very different in a small RFA compared to a large open grouping, but being 10 percent of the total amount requested away is more equivalent.[49]

In all of the models that use this instrument, I always include controls for the quality of the applications as judged by peer review. This is because lower-quality applications will, by construction, always be "further" from the top of the funding group. The goal of this instrument is to leverage variation in awards among applications that received equal marks by their peer reviewers. However, there are not any straightforward ways of providing strong empirical evidence that the exclusion restriction holds in this case. So some caution should be taken when interpreting the absolute magnitude of these coefficients and implied effects.

I estimate variants of equation (9), a two-stage least squares (2SLS) equation determining whether an applicant receives a grant ($\mathbf{1}\{Win_{jk}\} = \{0,1\}$), as well as the postdecision publication output of that same applicant (y_{jk}):

$$\mathbf{1}\{Win_{jk}\} = a_k + \delta z_{jk} + f_1(score_{jk}, rank_{jk}, \mathbf{X}_j) + v_{jk}, \tag{9a}$$

$$y_{jk} = \alpha_k + \beta \mathbf{1}\{Win_{jk}\} + f_2(score_{jk}, rank_{jk}, \mathbf{X}_j) + \nu_{jk}, \tag{9b}$$

where the score and rank variables capture the quality of the application based on its peer review score and ordinal ranking within its funding group, $\mathbf{X}_j$ is a vector of application/applicant-specific characteristics,[50] z_{jk} is the instrumental variable, and

[49] I use the funds requested, and not those actually awarded, because realized award amounts are obviously influenced by NIH preferences and may reflect strategic withholding of awards from better-ranked applications to fund more-preferred but lower-ranked applications.

[50] These other covariates include: funds requested by the focal application; whether the project is flagged as involving animals, humans, or children; whether the applicant is a "new investigator" per NIH's definition; whether the applicant has a PhD, MD, or other degree; whether the applicant has ever received an NIH research grant; the

the error terms are denoted by v_{jk} and ν_{jk}. In the simplest specifications, a_k and α_k represent institute-year fixed effects intended to capture secular changes in the use of open and RFA mechanisms.[51] In more stringent specifications, these fixed effects will be at the level of funding groups, which will only identify the productivity effect based on differences between winners and losers in the same RFA or open competitions.

I allow the focal productivity parameter, β, to vary by RFA and open awards, interacting z_{jk} with an RFA/open dummy to generate the necessary identifying variation. To handle the skewed nature of the outcomes while retaining zeros, I transform dependent variables using the inverse hyperbolic sine transformation and then transform the coefficients to report approximate semielasticities.[52]

To minimize parametric assumptions about the role of application quality (score and rank) and other covariates ($\mathbf{X}$), I use the *LASSO*-based variable selection and estimation routine described by Chernozhukov, Hansen, and Spindler (2015).[53] I separate these covariates into three categories: (i) the quality variables score and rank, (ii) a set of "project" $\mathbf{X}$ variables specific to the application, and (iii) a set of "people" $\mathbf{X}$ variables specific to the applicant. Within each of these categories, I generate fourth-order polynomial expansions of the variables and interact all of the linear terms. I then vary the degree to which the project and people controls and the funding group fixed effects are included in the model to investigate whether any differences between RFA and open grant productivity is driven by certain selection patterns.

C. *Summary Statistics and Results*

I evaluate outcomes for scientists who applied to the NIH between 2000 and 2006 and receive a peer review score. The sample includes only funding groups that award at least one grant and receive at least five applications, which removes the smallest 1 percent of groups. The final sample includes a set of 211 open funding groups, 334 RFAs, and 22,856 scientists who submitted 34,437 applications, about 15 percent of which went to RFAs. Online Appendix F includes a table with detailed summary statistics for this sample.

Figure 7 graphically depicts the first-stage relationship after controlling for the LASSO selected application quality variables and institute-year fixed effects. The instrument is expressed in terms of standard deviations. The relationship suggests that in both RFAs and open groupings, being one standard deviation further from the best-ranked application in the group leads to a roughly 20 percentage point decline

years since the applicant's first NIH grant of any kind and R01 grant; and the count of the applicant's publications including interactions for those as first/last author, the JIF weights, and whether the publication is flagged as receiving support from the US government. In the RFA-only similarity analyses, I also include a vector of preapplication publication similarity scores.

[51] The data is right-censored, with only half of the sample observed for the full five years postdecision. So if certain institutes relied on RFAs more in later years, then RFAs would appear less productive by construction.

[52] See Bellemare and Wichman (forthcoming) for more.

[53] These methods are implemented using the authors' lassopack Stata module.

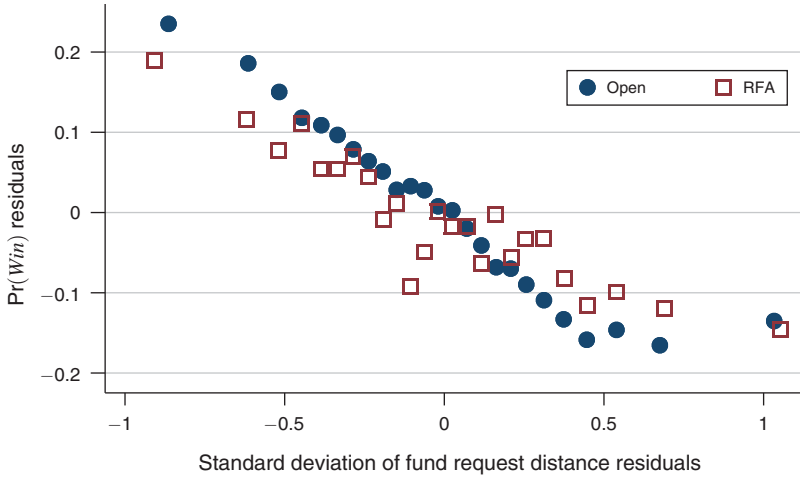

FIGURE 7. FIRST-STAGE RELATIONSHIP—RESIDUALS

Notes: This figure plots win probabilities and the funding instrument for each application, conditioning out LASSO-selected application quality and institute-year fixed effects. The *F*-statistics are 226.7 and 51.3 for open and RFA applications, respectively.

in funding probability. The budget constraint appears to gradually bind in a linear fashion, which motivates the choice of not transforming the instrument any further.[54]

Table 3 presents the results of estimating equation (9). Column 1 is a simple OLS version of the regression with only the institute-year fixed effects. This regression suggests that on average, RFA grants are roughly three times as productive as open grants. This difference could be attributed to higher quality applications being submitted to RFAs, more productive scientists—conditional on their peer review scores—selecting into RFAs, or the NIH selecting more promising areas of science for RFAs.

To hold application quality fixed and identify estimates closer to the causal effect of the grants themselves, column 2 uses the IV and still finds RFAs to be roughly three times as productive as open grants. But by relying on the IV, I can be more confident that this is not due to differences in the quality of marginal RFA and open submissions and that the absolute magnitudes identified reflect the impact of the grants themselves. The absolute productivity levels are larger than the OLS estimates, which may be due to preferences of the NIH for project- or people-specific characteristics aside from publication potential or some other feature unique to the compliers of this instrument (i.e., the local average treatment effect per this IV may be smaller than the true average treatment effect).

[54] The first-stage relationship plotted in Figure 7 corresponds to the IV model in Table 3, column 2. In online Appendix F, I present regression tables of the first-stage relationships for the five IV models in Table 3, columns 2–5. The results are very similar across the models, with different controls for projects, people, and other fixed effects. In all cases, for both RFA and open groupings, a one standard deviation increase in the instrument leads to a roughly 17 to 24 percentage point decline in funding probability, conditional on the application quality metrics. And the *F*-statistics of the instruments are all on the order of 100 to 200 for open applications and 50 for RFA applications.

Table 3—Grant Productivity—Publication Rates

	IHS(*Publication Count*$_{jk}$)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\{Win, open_{jk}\}$	0.0578	0.224	0.250	0.268	0.0657	0.129
	(0.0206)	(0.0752)	(0.0781)	(0.0767)	(0.0565)	(0.0477)
$\mathbf{1}\{Win, RFA_{jk}\}$	0.134	0.620	0.628	0.451	0.0938	0.133
	(0.0362)	(0.142)	(0.144)	(0.163)	(0.0599)	(0.0452)
Semielasticity *open*	0.059	0.247	0.280	0.304	0.066	0.137
Semielasticity *RFA*	0.142	0.841	0.854	0.550	0.096	0.141
p-value difference	0.07	<0.01	<0.01	0.24	0.63	0.93
Observations	34,437	34,437	34,437	34,437	34,437	34,437
IV		Y	Y	Y	Y	Y
F-statistic		122.7	121.1	115.7	79.7	54.3
Project **X**			Y			Y
People **X**				Y		Y
Funding group fixed effects					Y	Y
LASSO var$_{sel/poss}$		5/9	14/31	27/256	5/9	34/287

Notes: This table reports 2SLS estimates from equation (9). The average applicant published 6.24 articles postdecision. Project and People **X** indicate whether covariates specific to the application (i.e., funds requested) and/or the applicant (i.e., publication history) are included. *LASSO var*$_{sel/poss}$ reports the number of *LASSO* selected and possible covariates. All regressions include institute-year fixed effects, except for columns 5 and 6, given the funding group fixed effects, which are within institute years. Standard errors clustered at funding groups are in parentheses.

Columns 3–6 investigate what may be leading to this productivity difference. First, column 3 introduces the vector of project covariates (i.e., budget size, whether the project involves human subjects) to remove some variation in the types of science proposed. The estimates are virtually unchanged. Column 4 includes only the people-specific covariates (i.e., publication and grant histories), and the productivity difference declines by nearly 60 percent, relatively speaking. This suggests that a large reason why RFAs appear more productive is that highly productive scientists, whose future productivity is not captured in the peer review score, select into these mechanisms.

Can the rest of the difference be attributed to the types of science targeted by RFAs? The inclusion of the funding group fixed effects in column 5 identifies RFA and open award semielasticities that are not significantly different from each other or zero. This indicates that a sizable portion of the RFA-open difference could plausibly be attributed to NIH program managers identifying research areas that have high publication potential. Although this model suggests grant receipt does not have a clear effect on publication counts, this does not imply that RFAs are ineffective. Rather, it suggests that RFAs may be able to identify research directions with high potential that scientists would not have taken otherwise.[55] And part of this effect may be due to scientists applying and losing but continuing to work on the targeted topic for some time.

[55] Furthermore, this is not necessarily at odds with the RFA endogeneity tests in Section IC, since the argument for RFA exogeneity hinges on the NIH not responding to shocks that scientists themselves were also responding to. And this result suggests the NIH may be able to respond to (future) shocks that scientists would not have responded to.

Table 4—Grant Productivity—Publication Similarity

	IHS(*Publication-RFA Similarity*$_{jk}$)		
	(1)	(2)	(3)
$\mathbf{1}\{Win, RFA_{jk}\}$	0.131	0.334	0.317
	(0.0328)	(0.166)	(0.136)
Semielasticity *RFA*	0.140	0.378	0.361
Observations	4,949	4,949	4,949
IV		Y	Y
F-statistic		57.5	58.2
Project, people **X**			Y
Funding group fixed effects	Y	Y	Y
pmra controls	Y	Y	Y
LASSO $var_{sel/poss}$	3/21	6/21	12/350

Notes: This table reports estimates from equation (9) using the *pmra*-based scientist-RFA similarity measure. Project and people **X** indicates whether covariates specific to the application (i.e., funds requested) and/or the applicant (i.e., publication history) are included. *LASSO* $var_{sel/poss}$ reports the number of *LASSO* selected and possible covariates. Standard errors clustered at funding groups are in parentheses.

In the fully saturated specification of column 6, the productivity difference is no longer statistically or economically meaningful. Both grants lead to about 14 percent more publications, and the *p*-value for the test of difference between the coefficients is 0.93.[56] This is rather intuitive since there is not any clear difference between the two grant mechanisms after removing variation in the types of applicants, applications, and scientific fields involved.[57]

Finally, I focus on RFAs and the question as to whether these grants actually generate new studies on the topics targeted. Table 4 presents estimates from another set of productivity equations using the *pmra*-based similarity outcome. I only use specifications that include both the funding group (= RFA) fixed effects and measures of the scientists' similarity to the RFA prior to applying. This is because there is some (unknowable) amount of variation in similarity scores across scientists and RFAs that is purely due to differences in how each RFA and publication was written and evaluated by the *pmra* algorithm. Always using these fixed effects and covariates prevents an in-depth investigation into selection patterns, but it ensures that the identifying variation in similarity scores is not mechanically related to differences in text usage.

Column 1 presents the OLS estimate, and columns 2–3 use the funding distance instrument. As before, the IV estimates are larger than the OLS estimate. In the IV specifications, regardless of whether the project and people covariates are included, the instrumented grants induce a roughly 37 percent increase in the similarity between scientists' publications and the RFA's objectives. In this sample, this

[56] For comparison, Jacob and Lefgren's (2011) headline finding, using a specification most closely aligned with Table 3, column 6, is a 7 percent increase in publications. The difference in our results can plausibly be explained by our alternative identification strategies.

[57] Results in online Appendix F using the JIF measure reveal a somewhat similar pattern of relative magnitudes between the RFA and open semielasticities. Although, the absolute magnitudes are closer to zero, and the people covariates alone can plausibly explain the small RFA-open quality difference.

magnitude corresponds to approximately one-third of a standard deviation in similarity scores, suggesting this effect is meaningful in a scientific sense.

Online Appendix F contains time-specific estimates of the instrumented grant effects. The publication rate effects appear to be concentrated in the first three years postdecision, with the estimates being indistinguishable from zero by the fifth year. This likely reflects the average four-year lifespan of the grants in this sample. For RFA grants, the impact on similarity appears to be almost entirely concentrated in the first year postdecision. Whether this is due to the losers becoming more similar to the winners (and the RFA) or the winners returning to their loser-like trajectory is difficult to discern without a second control group. But the raw data can shed some light: the average *pmra* similarity scores for both winners and losers of RFAs decline by about 30 percent from the first to fifth year post decision. This suggests that the latter story of winners returning to their original trajectories once their grant expires is most likely.

D. *Back-of-the-Envelope Cost Effectiveness*

To recap, the analyses in Section III indicated that RFA awards are approximately 16 to 20 percent larger, which corresponds to roughly an additional $200,000 over the lifespan of the average grant. A conservative use of the estimates from Table 3 implies that RFAs can produce at least 1.8 more publications per grant than the open channels (3.1 versus 1.3).[58] Since the average lifetime size of an R01 in my data is about $865,000, these effects imply that the average cost per publication is about $344,000 for RFA grants and $665,000 for open grants. RFAs appear to be significantly more cost effective at generating new publications.

But this does not immediately imply that the RFA program should be expanded; two caveats are in order. First, this productivity difference appears to be driven by the different types of scientists or projects that RFAs fund. So it is not clear how scaling the RFA programs beyond current levels would change this composition. If this composition effect is eliminated (as in Table 3, column 6), and the RFA premium remains, then policymakers will face a more difficult trade-off between the two mechanisms.

Second, the time trends of similarity effects illustrated in online Appendix F indicate that scientists' focus on the RFA-targeted topics does not appear to be persistent, so only a fraction of these additional publications are directly in line with the NIH's objectives. It appears that larger or more sustained funding would be necessary to make changes to the long-run direction of scientists.

More generally, I cannot extend these results to considerations of "moonshot"-style policies such as Nixon's "War on Cancer" that most certainly have the potential to create larger changes in the general equilibrium. While I cannot speak to the absolute welfare impacts of these grants (i.e., how large should the total RFA plus open budget be?), these estimates still provide some guidance to policymakers debating the merits of using RFAs at current levels to steer scientists toward topics they value.

[58] This is based on taking the average difference in the RFA and open grant semielasticities across Table 3, columns 2–6.

V. Conclusion

The preceding trio of analyses finds the following: The scientists most responsive to RFAs are already operating relatively close to the topics that are targeted. The general costs of changing directions implied by scientists' RFA entry decisions appear very large. And these adjustment costs can plausibly explain the RFA premium, whereby RFA grants tend to be larger that their open counterparts. Irrespective of content, RFA grants lead to more publications than open grants, but this difference appears to be due to the different types of science and scientists that RFAs fund. Finally, RFAs generate new publications that are in line with the original objectives, but recipients appear to eventually revert to their original research directions.

Focusing on the large adjustment costs, the welfare implications of this constraint are not immediately obvious. For instance, although it may appear that the inelasticity I identify could prevent scientists from choosing an efficient direction from society's view, they may in fact be a source of diversity that counterbalances other distortions arising from market dynamics (Acemoglu 2012) or racing incentives (Stephan 1996). Still, the magnitudes suggest that these costs are a first-order force in determining the allocation of scientists.

In the context of research funding mechanisms, these results are relevant to the ongoing "people versus projects" debate on the optimal structure of grants.[59] Traditionally, and at the NIH especially, research grants are awarded for "projects." However, based on a growing body of theoretical (Manso 2011) and empirical evidence (Azoulay, Graff Zivin, and Manso 2011), calls have grown for more flexible funding arrangements that leave more discretion to the scientist. A frequent critique of project-centered arrangements is that they reduce incentives to be "creative."

The large switching costs I identify provide another potential limitation of project-based funding regimes: if researchers can only successfully compete for project-based grants by demonstrating their ability to produce publications, then they will not propose projects that require large adjustments, because they would appear to be unproductive in the short term (as funds would be spent on these switching costs, e.g., purchasing new tools, and not directly on production, e.g., labor). However, I am unable to speak directly to any of the mechanisms underlying the apparent reluctance of scientists to change their work. Disentangling forces such as information barriers, risk preferences, and tangible costs would be useful for informing policies.

Since Oi (1962) first emphasized the notion of labor as a quasi-fixed factor, "the hypothesis that employment adjusts slowly and with a speed that is inversely related to skill, [has] entered the central corpus of economic knowledge" (Hamermesh 1990, 94). Likewise, since at least Schmookler (1966), demand's ability to "pull" invention has been appreciated. Here, instead of examining the costs of and motivations for adjusting the *rate* of production, as has most Oi- and Schmookler-inspired work, I use RFAs to understand the costs of incentivizing adjustments to the *direction* of work for

[59] See, for example, https://goo.gl/TEQ7W8, accessed July 12, 2017.

one of the most skilled sets of the economy, biomedical scientists. My results show that at least in this setting, these costs are likely of first-order relevance.

Understanding this elasticity is essential for policymakers and managers alike since the vast majority of scientists at public and nonprofit institutions choose their own pursuits with minimal oversight. This system has arisen for good reason: uninformed funders are willing to relinquish control to leverage scientists' private information (Aghion, Dewatripont, and Stein 2008). However, it is not clear that the allocation of funds decided by scientists themselves will be socially optimal (Dasgupta and David 1994).[60] To resolve such concerns, Aghion, Dewatripont, and Stein (2008) suggests the use of mechanisms analogous to the RFAs I study here.

An obvious next question is how these costs influence the optimal allocation of research funds. A few theoretical models study this question (e.g., Lichtenberg 2001), but none reckon with the type of costs emphasized here. Future work on this topic should be worthwhile.

REFERENCES

Acemoglu, Daron. 2002. "Directed Technical Change." *Review of Economic Studies* 69 (4): 781–809.

Acemoglu, Daron. 2012. "Diversity and Technological Progress." In *The Rate and Direction of Inventive Activity Revisited*, edited by Josh Lerner and Scott Stern, 319–56. Chicago: University of Chicago Press.

Aghion, Philippe, Antoine Dechezleprêtre, David Hémous, Ralf Martin, and John Van Reenen. 2016. "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry." *Journal of Political Economy* 124 (1): 1–51.

Aghion, Philippe, Mathias Dewatripont, and Jeremy C. Stein. 2008. "Academic Freedom, Private-Sector Focus, and the Process of Innovation." *RAND Journal of Economics* 39 (3): 617–35.

Aguirregabiria, Victor, and Pedro Mira. 2007. "Sequential Estimation of Dynamic Discrete Games." *Econometrica* 75 (1): 1–53.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, edited by Universities-National Bureau Committee for Economic Research and Committee on Economic Growth of the Social Science Research Council, 609–26. Princeton: Princeton University Press.

Azoulay, Pierre, Christian Fons-Rosen, and Joshua S. Graff Zivin. 2019. "Does Science Advance One Funeral at a Time?" *American Economic Review* 109 (8): 2889–2920.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat. 2019. "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules." *Review of Economic Studies* 86 (1): 117–52.

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. 2011. "Incentives and Creativity: Evidence from the Academic Life Sciences." *RAND Journal of Economics* 42 (3): 527–54.

Bajari, Patrick, Han Hong, John Krainer, and Denis Nekipelov. 2010. "Estimating Static Models of Strategic Interactions." *Journal of Business and Economic Statistics* 28 (4): 469–82.

Bellemare, Marc F., and Casey J. Wichman. Forthcoming. "Elasticities and the Inverse Hyperbolic Sine Transformation." *Oxford Bulletin of Economics and Statistics*.

Borjas, George J., and Kirk B. Doran. 2015. "Cognitive Mobility: Labor Market Responses to Supply Shocks in the Space of Ideas." *Journal of Labor Economics* 33 (S1, Pt. 2): S109–45.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler. 2015. "Post-selection and Post-regularization Inference in Linear Models with Many Controls and Instruments." *American Economic Review* 105 (5): 486–90.

Crane, Diana. 1969. "Social Structure in a Group of Scientists: A Test of the "Invisible College" Hypothesis." *American Sociological Review* 34 (3): 335–52.

[60] The "non-congruence between private and social rankings of final outcomes creates fundamental grounds for suspecting that the research portfolio that would be, in effect, selected, for society by the self-governing community of scientists will be an inefficient one" (Dasgupta and David 1994, 506).

Dasgupta, Partha, and Paul A. David. 1994. "Toward a New Economics of Science." *Research Policy* 23 (5): 487–521.

Hamermesh, Daniel S. 1990. "Aggregate Employment Dynamics and Lumpy Adjustment Costs." *Carnegie-Rochester Conference Series on Public Policy* 33: 93–129.

Jacob, Brian A., and Lars Lefgren. 2011. "The Impact of Research Grant Funding on Scientific Productivity." *Journal of Public Economics* 95 (9–10): 1168–77.

Krasnokutskaya, Elena, and Katja Seim. 2011. "Bid Preference Programs and Participation in Highway Procurement Auctions." *American Economic Review* 101 (6): 2653–86.

Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lerchenmueller, Marc J., and Olav Sorenson. 2016. "Author Disambiguation in PubMed: Evidence on the Precision and Recall of Author-ity among NIH-Funded Scientists." *PloS One* 11 (7): e0158731. https://doi.org/10.1371/journal.pone.0158731.

Lichtenberg, Frank R. 2001. "The Allocation of Publicly Funded Biomedical Research." In *Medical Care Output and Productivity*, edited by David M. Cutler and Ernst R. Berndt, 565–90. Chicago: University of Chicago Press.

Lin, Jimmy, and W. John Wilbur. 2007. "PubMed Related Articles: A Probabilistic Topic-Based Model for Content Similarity." *BMC Bioinformatics* 8: 423 (2007). https://doi.org/10.1186/1471-2105-8-423.

Manso, Gustavo. 2011. "Motivating Innovation." *Journal of Finance* 66 (5): 1823–60.

Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish Émigrés and US Invention." *American Economic Review* 104 (10): 3222–55.

NBER. 1962. *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton: Princeton University Press.

Nelson, Richard R. 1959. "The Simple Economics of Basic Scientific Research." *Journal of Political Economy* 67 (3): 297–306.

Oi, Walter Y. 1962. "Labor as a Quasi-fixed Factor." *Journal of Political Economy* 70 (6): 538–55.

Schmookler, Jacob. 1966. *Invention and Economic Growth*. Cambridge: Harvard University Press.

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

Stern, Scott. 2004. "Do Scientists Pay to Be Scientists?" *Management Science* 50 (6): 709–853.

Torvik, Vetle I., and Neil R. Smalheiser. 2009. "Author Name Disambiguation in MEDLINE." *Association for Computing Machinery Transactions on Knowledge Discovery from Data* 3 (3) Article 11 (July 2009). https://doi.org/10.1145/1552303.1552304.

Williams, Heidi L. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." *Journal of Political Economy* 121 (1): 1–27.