The Elasticity of Science

Kyle Myers*

March 8, 2018

Abstract

This paper provides the first estimates of the degree to which scientists are willing to change their direction in exchange for resources. Novel data based on funding at the National Institutes of Health is used to estimate an entry model that accounts for strategic interactions. The results indicate that inducing a scientist to change their direction by 1 standard deviation, a qualitatively small change in their science, requires a four-fold increase in funds, an additional $1 million per year. Furthermore, this responsiveness is very heterogenous, competitive effects are important, and targeted funds do stimulate new science in policymakers' intended direction.

JEL Codes: H50, I23, O31, O33, O38

*National Bureau of Economic Research; 1050 Massachusetts Ave., Cambridge, MA 02138; 617-868-3900; kmyers@nber.org; current version available at kylemyers.org. 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, who welcomed me to the NIH, 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.

1 Introduction

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 the knowledge that drives economic growth (Stephan 1996), it is essential to understand how costly it is to incentivize changes in the direction of their work - the elasticity of science. This paper provides the first estimates of this elasticity based on a novel administrative dataset of targeted funding at the world's largest scientific agency, the U.S. National Institutes of Health (NIH).

While the rationale for the public support of science has been appreciated for some time (cf. Nelson 1959, Arrow 1962), it has largely been assumed that when governments direct funds to particular pursuits, scientists will simply follow. But how costly, for example, would it be to incentivize a scientist studying one disease to pursue another? There is no clear evidence as to the magnitude of these costs, or how they vary across scientists. Quantifying this sort of responsiveness is important because, as emphasized by models of directed technical change, the optimal research policy will depend on a market's directional elasticity (Acemoglu 2002).[1]

Measurement and identification challenges have limited progress on this topic to date, despite it being more than 50 years since economists set their sights on understanding the rate *and direction* of inventive activity (Nelson et al. 1962). The main difficulties are that new opportunities often arise endogenously and scientists can self-select their pursuits. Accounting for these factors econometrically is compounded by the difficulties of observing and quantifying an individual scientist's "direction".

The data and methods in this paper overcome these challenges by leveraging exogenous variation generated by a series of targeted grant competitions at the NIH and a validated algorithm that identifies changes in scientists' directions. There are three key findings: first, it takes relatively large changes in funding amounts to incentivize small changes in direction; second, this inelasticity creates a wedge in payoffs where grants requiring redirection are substantially larger in expected value; and finally, the observed redirections are real - grant receipt has a causal impact on both the rate and direction of scientists' publications.

[1] In these models, a key parameter that determines the benefits to pursuing new directions (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 based on the presence of within-factor spillovers and/or across-factor switching costs. This paper is focused on the latter.

At the NIH, the majority of grants are awarded through "investigator-initiated" competitions, which cater to all types of biomedical research. But routinely, the NIH sets aside funds (roughly $2 million at a time) for one-time competitions, which solicit proposals on specific diseases, populations, and/or methodologies. These are termed Requests For Applications (RFAs). Consider this remark by NIH Director Francis Collins:

> "[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"[2]

It is implied that by allocating funds to certain topics, the NIH believes it can effectively steer researchers to those topics, just as a supply or demand shock might push or pull scientists in a particular direction. However, there is no evidence on how costly it is to bring "new brains" into a field, and whether or not these "new brains" can contribute.[3]

The RFA setting is well suited to tackle these questions. 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 of modern biomedical research, this is a useful and much more practical alternative to a traditional randomized experiment. Furthermore, in the process of identifying the focal parameter, I test whether or not the funds directed by the NIH are actually used as intended, an assumption at the core of the Nelson-Arrow paradigm.[4]

Specifically, I identify how costly it is to incentivize directional adjustments by comparing the implied weight that scientists place on the *scientific similarity* of their prior work and a new project, relative to the amount of funds they could expect when pursuing that project. Figure 1 makes clear the importance of similarity and award size in scientists' decisions. Panel 1a shows that scientists are much more likely to enter an RFA when the objectives are more similar to their prior work. Panel 1b shows that scientists prefer to enter RFAs in which more funds are made available. Both results are intuitive, but this is the first clear illustration of these facts.

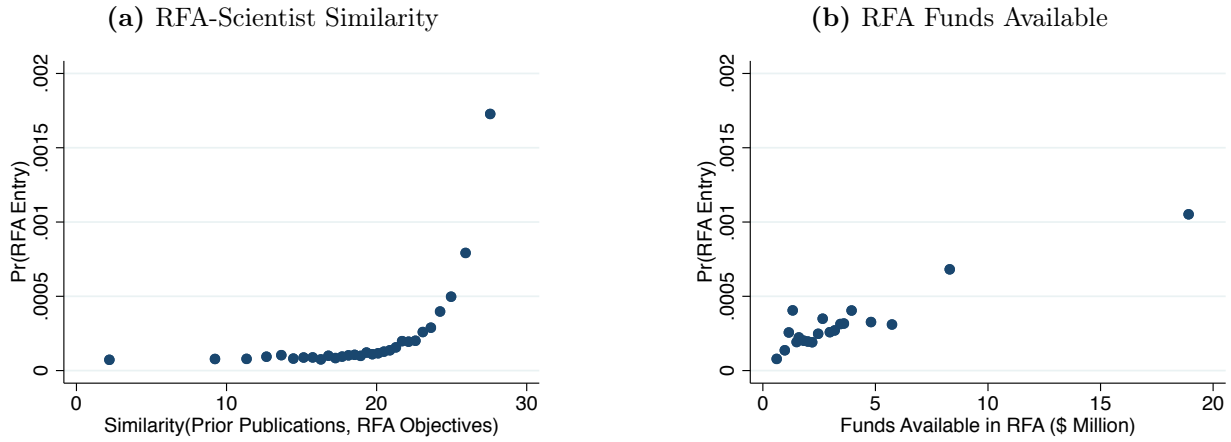
Figure 1 also begins to illustrate the elasticity of science, which, in this context, is the relative

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

[3]Certainly, economists have appreciated the adjustment costs of changing input levels (cf. Oi 1962). But only recently has work begun to identify directional adjustment costs of R&D activities, most of which has focused on private firms in the energy sector (i.e., Aghion et al. 2016).

[4]Recall, Nelson and Arrow's conclusion was not to simply fund "science" broadly construed, 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 (i.e., RFAs).

Figure 1: Probability of RFA Entry per Similarity and Funding



Note : Binned scatterplots of scientists' entry probabilities per: Panel (a), similarity of scientists' prior publications to the research objectives of the RFA (larger scores indicate greater similarity); Panel (b), the amount of funds made available in the RFA.

value scientists place on similarity (Panel 1a) versus expected award (Panel 1b). However, simply comparing the relative slopes implied by Figure 1 would ignore competitive effects, and assume that scientists entering an RFA would not have pursued those topics otherwise. The empirical analyses in this paper are all centered on accounting for the first point and verifying the second.

The major identification concern is that if the creation of RFAs is correlated with changes to the supply or demand for the topics they target (e.g., a disease outbreak), then one cannot discern whether scientists are responding purely to the NIH's funds or to these (unobservable) events. But throughout the paper I show evidence that there is no economically meaningful correlation between RFA generation and other underlying trends in the science targeted. Notably, there are no pre-trends that would indicate that scientists are increasingly interested in targeted topics prior to each RFA.

My empirics use roughly a dozen years of administrative data, covering about 400 RFAs, and publication histories for approximately 140,000 biomedical scientists. I conduct three core analyses. In the first, I leverage variation within and across RFAs to estimate an entry model in the spirit of Bajari et al. (2010), accounting for the endogeneity of strategic interactions. Relating the marginal value of scientific similarity to the marginal value of RFA funds, while holding competitive expectations fixed, identifies an elasticity of 0.2. Not controlling for competitive effects significantly underestimates the elasticity. The qualitative meaning of this estimate is discussed in depth, and by all measures it indicates that individual scientists

are inelastic. Still, the base of scientists is quite large. Exploring the role of this parameter in the aggregate, simple counterfactuals suggest that if scientists were 10% more (less) elastic, the marginal cost of incentivizing RFA applications would decrease (increase) by roughly 5%. However, competition for funds appears to place a lower bound on how cheap applications could become.

In the second analyses, I examine to what extent this force gives rise to an "RFA premium." I follow the argument that if RFAs induce scientists to adjust their direction, and these adjustments are costly, then the expected value of RFAs could be larger than the grant mechanism where scientists choose their own pursuits. Controlling for selection bias with important covariates (i.e., individual-fixed effects, quality scores), I show that, relative to the investigator-initiated grants, RFAs induce scientists to change their direction by 10% and the RFA premium is 65% of the average grant size. The relative magnitudes of these effects indicate that virtually all of the RFA premium can be attributed to the inelasticity of science.

The first two analyses make use of data contained within scientists' *applications*, which may not capture the realized changes in their work. In the final analyses, I test whether these intentions are actually pursued. Using a regression discontinuity design made possible by the rank-order funding mechanism, I show that this is indeed the case. Conditional on applying to an RFA, marginal winners publish more and undertake relatively larger redirections towards the RFA than marginal losers.

An elasticity of 0.2 implies that scientists are indifferent between pursuing a new project that is 1 s.d. more similar than average and a four-fold increase in expected funding, an additional $1 million per year for about four years. And a 1 s.d. difference in similarity is by no means dramatic. It is equivalent to a virologist switching between testing vaccines for two closely related RNA viruses, or, more loosely speaking, an economist switching between studying the supply or demand for pharmaceutical R&D.[5] It is rather surprising then that virtually zero funds from any major scientific funding agency in the world explicitly subsidize field-to-field transitions.[6] This paper makes clear that the costs of redirecting scientists should have first-order implications for the optimal allocation of research.

My results complement a long line of sociological studies that emphasize the role of non-pecuniary forces in science (i.e., Crane 1969, Kuhn 1970, Merton 1973). More recently, Stern

[5] An additional $1 million per year would more than quadruple the average grant size. Section 2.4 graphically depicts the similarity distribution with these virologist and economist examples.

[6] 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 pre-specify the destination field and are rarely used.

(2004) identifies a large wage premium associated with the right to publish, indicating scientists have strong preferences, at least when choosing a career. Closely related is Azoulay, Fons-Rosen and Graff Zivin (2016) who study how scientists choose which topics pursue, as in this paper. 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. I build on this work in identifying how pecuniary incentives can shape a scientist's direction given the cumulative role of preferences, perceived barriers, or any other constraint.

2 Setting & Data

2.1 NIH Overview

Broadly speaking, the NIH's main objective is to award roughly $28 billion annually 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% in the 1980s to roughly 30% 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, with awards prominently displayed on individuals' CVs. Scientists take great care to stay abreast of the NIH and commit extensive time to grant pursuits. The application decision is not taken lightly.

The grants I study provide funds for "projects." Applications for these grants propose a self-contained research idea. Funds awarded may be used to purchase 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% of all grant awards (70% of funds). These are "award[s] made to support a discrete, specified, circumscribed project," and on average award $285,000 in their first year with the average award lasting 4.2 years.

When seeking funding at the NIH, scientists have two major options: RFAs and the "investigator-

[7] Awards have two distinct 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 throughout. Robustness tests in the Appendix show the main results hold when examining total costs. Unless otherwise specified, the remainder of the paper refers to direct costs when discussing awards.

initiated," or what I refer to as "open," grants. The Appendix 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, which is continuously available, scientists may 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.

The key differences between RFA and open grant competitions are as follows.[8] Open competitions are roughly 20 times larger than RFAs in terms of the total funds and number of awards to be awarded. 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 (of which there are about 175 standing peer review panels). The breadth of science is much larger in open competitions. Most notably, open applications have a win probability of 16.3% with an average award size of $275,000, whereas RFA applications have a win probability of 19.4% with an average award size of $339,000. First-time RFA applications are both more likely to win, and conditional on winning, are larger. Section 4 focuses on this apparent wedge to identify how much of it is truly due to the costs facing scientists when adjusting the direction of their work.

2.2 Data Sources

NIH Applications & 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 contains the following: application and applicant identifiers; peer review grouping and score; funding decision and award size; Institute; fiscal year. For applications submitted on or after 2006, the data also contains the abstract and title of the proposed research project for both funded and non-funded applications.

RFA Details—: The research objectives, funds allotted, and timing of each RFA announced between fiscal years 2002 and 2009 were scraped from the NIH announcement website and manually reviewed to ensure accuracy.[9] A total of 1,125 RFAs were scraped. I restrict the sample of RFAs to include only those that solicit R01 grants (686), were not released as a part of the American Reinvestment and Recovery Act (678), and do not request "renewal"

[8] In both cases, there are 20-25 unique competitions of either type available at any given time.
[9] Available at https://goo.gl/LuaBOQ, accessed July 12, 2017.

grant applications (537).[10] 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 2006-2009 because this approach requires the application abstract data, which is only available in applications 2006 onward.[11]

Publications—: Each scientist's publication record (regardless of funding) prior to 2009 was 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) develop 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).

2.3 RFA Generation & 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.[12] This is relevant to two aspects of this study: (1) identification: are scientists responding to the NIH's funds or other correlated events; and (2) 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. In order 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.[13]

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

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

[12]This overview is based on interviews with NIH staff. The Appendix outlines these major motivations in further detail.

[13]Using the count of abstracts associated with specific MeSH terms follows the same logic as prior work which proxies for the scientific importance of particular genes with the number of publications related to

The Appendix provides further details about sample construction and displays the cross-sectional distribution of treated and control terms. On average, RFA treated terms occur in publications and applications at about a 30% larger rate than the control terms. RFAs tend to focus on topics more popular than average.[14] 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 "pre-trends".[15] To empirically test for these pre-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:

$$\mathbb{E}[N_{mt}|\delta_m, \gamma_t] = \exp\Big(\sum_{\tau=t-\underline{s}}^{t+\overline{s}} \beta_\tau \times \mathbf{1}\{\text{RFA}_{m\tau}\} + \delta_m + \gamma_t \Big), \tag{1}$$

where $\mathbb{E}$ is the expectations operator, δ_m and γ_t are MeSH- and time-fixed effects, respectively, and $\mathbf{1}\{\text{RFA}_{mt}\}$ equals one when MeSH term m is associated with at least one RFA at time t.[16] 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 $\overline{s}$ periods after t. In all models, β_τ is estimated for four years prior to t and two years post t, which corresponds to ($\underline{s}$, $\overline{s}$) = (4,2) for the PubMed data given annual observations, and ($\underline{s}$, $\overline{s}$) = (12,6) for the NIH data given three observations per year.[17] 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 pre-trends.

Focusing on the NIH data, in the time period when the RFAs occur, there is a sharp increase

each gene (e.g., Williams 2013).

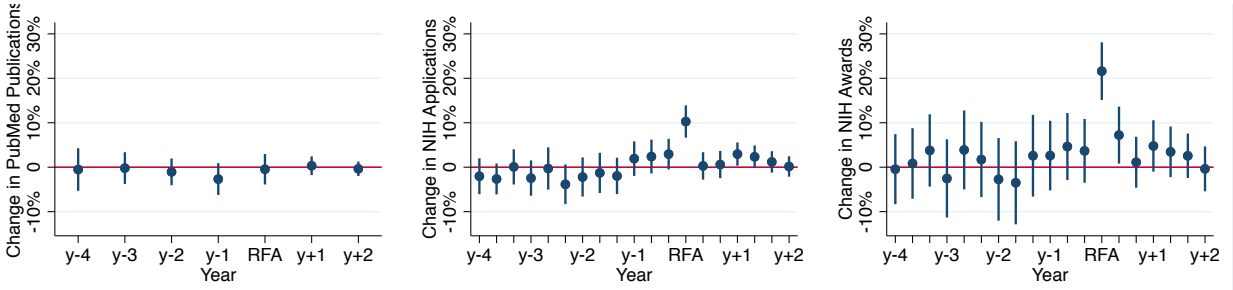
[14] What this means for generalizability will depend on whether the costs of incentivizing scientists to pursue topics varies across this distribution. On one hand, the costs of pursuing low-rate subjects may be higher if the low rate is indicative of the low net value associated with those topics. 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.

[15] 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.

[16] 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.

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

Figure 2: Event Study of RFA-Targeted Subjects relative to Controls



(a) PubMed Publications **(b)** NIH Applications **(c)** NIH Awards

Note : Plots the β_τ coefficients estimated using Eq. 1. 95% confidence intervals (based on standard errors clustered at the MeSH term level) are plotted as bars.

in applications ($\beta_{\text{RFA}} = 10\%$) and awards ($\beta_{\text{RFA}} = 20\%$) associated with the treated MeSH terms.[18] 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 that the response observed is due predominately to RFAs specifically, and not any spillovers.[19] These event studies support the main identification assumption of this paper: scientists value RFAs in and of themselves.

2.4 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, or 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

[18]The lack of a treatment effect in the PubMed data is likely due to both the much larger scale of publications compared compared to applications (avg. MeSH rate of 322 compared to 4), and the slow, variable process by which NIH awards eventually give rise to new publications. The analysis of Section 5 focuses more specifically on this question of whether RFA awards do in fact lead to new science.

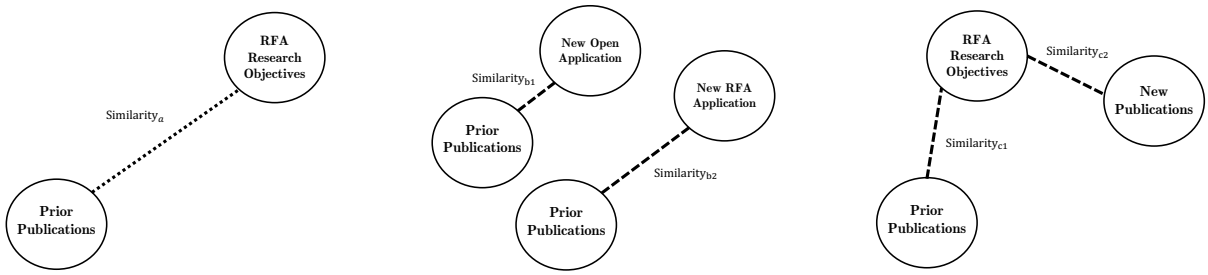
[19]For instance, if RFAs were a strong signal of future funding opportunities at the NIH, then we would have expected to observe significant post-trends.

Figure 3: Scientific Similarity & the Empirical Analyses

(a) Section 3, Entry Model (Within- & Across RFAs)

(b) Section 4, RFA Premium (RFA vs. Open)

(c) Section 5, Fuzzy RD (Within-RFAs)



and Wilbur (2007).[20] The Appendix outlines the algorithm in more detail and provides examples for 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.[21]

Figure 3 outlines how this metric will be used and provides a roadmap of the empirical analyses. First, in Section 3, I estimate RFA entry probability as a function of each scientist's similarity to the RFA (Similarity$_a$), and relate this to the role of RFA funding amounts to identify the elasticity. In Section 4, I test whether RFAs induce scientists to undertake projects that are less scientifically similar than they otherwise would have pursued in the open mechanism (i.e., Similarity$_{b1} \gtreqless$ Similarity$_{b2}$), and how this influences payoffs. Finally, in Section 5, I use a fuzzy regression discontinuity to identify whether grant receipt redirects scientists to topics more closely related to the objectives of the RFA (i.e., is (Similarity$_{c2}$ - Similarity$_{c1}$) larger for marginal winners?).

pmra 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 knowledge generated. 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.[22] The logic is that if publications define the boundary of each scientist's knowledge, the maximum

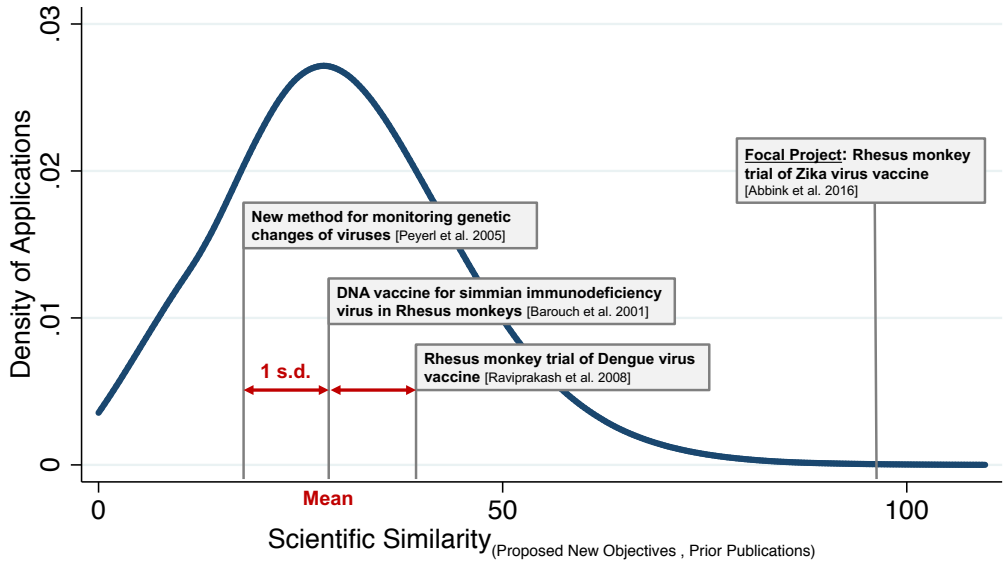
[20] *pmra* is the algorithm underlying the "Similar articles" feature of PubMed and has become a benchmark within the field of bioinformatics for measuring similarity.

[21] This algorithm has been used in similar work by Azoulay, Fons-Rosen and Graff Zivin (2016). 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)

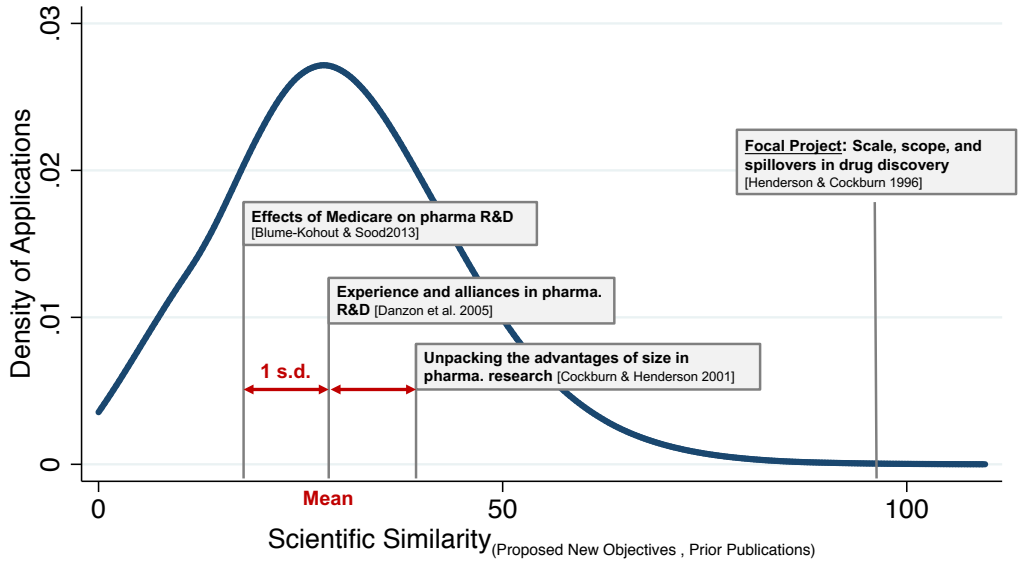
[22] For any applications listed with co-principal investigators (i.e., two or more scientists are project leads) I use the maximum of the two scientists' scores to be as conservative as possible.

Figure 4: Prior Publication - New Application Similarity Distribution & Examples

(a) Biomedicine: Virology



(b) Economics: Pharmaceutical R&D



Note : Article citations and further discussion in the Appendix.

captures the shortest proximity between the new science and the scientist.[23]

In order to get a qualitative sense of the magnitudes involved, Figure 4 plots the distribution of similarity scores for each NIH application (Similarity_b from Figure 3). Overlaid at relevant sample moments are publications from biomedicine and economics. The examples take a focal publication, e.g., a scientist's most recent article, and ask what would that scientist's next project look like given the NIH sample moments.[24] The mean indicates what the average "next project" would look like, and is flanked by projects at $+/-$ 1 standard deviation more/less similar. The figures make clear that virologists developing vaccines for viral diseases and economists studying pharmaceutical R&D tend to keep doing so. Previewing the focal result, the analyses below indicate that these example scientists (who wrote the "focal" papers) are indifferent between pursuing the mean next-project instead of the + 1 s.d. next-project and a fourfold increase in NIH grant funding (in expectation).

3 Entry Model & Elasticity Estimates

The focal elasticity is the relative change in the similarity of a scientist's next project that can be induced by a change in expected funding. Ideally, to identify this parameter, a random menu of grants that vary in terms of the their similarity and size would be offered to scientists. If allowed to choose from this menu, scientists' revealed preferences would identify the point of indifference, which quantifies the dollar amount necessary to induce some level of redirection. Conducting such an experiment is practically prohibitive given the real costs of biomedical research, but RFAs are a setting where scientists must make this precise tradeoff. The NIH creates a plausibly exogenous menu of RFAs and scientists are forced to ask themselves which, if any, of these should they pursue. Here, the elasticity can be identified by examining scientists' entry decisions and comparing the implied value they place on similarity and fund availability. However, the nature of competition - and therefore the manner in which the marginal value of these variables should be related - is not immediately obvious. The next subsections outline a simple entry model to motivate the following analyses.

[23]Certainly, the density of one's knowledge may vary within these boundaries, but using the maximum ensures the measure captures only variation in similarity in information rater than its depth (or quality). Robustness checks reported in the Appendix using other transformations (i.e., mean, median), restricting the publication set to articles published within the last 5 years, or controlling for the number of publications show no qualitative differences in the results.

[24]The focal publications are plotted at the average *pmra* score for documents scored against itself, noting that *pmra* scores are likelihoods, not percentages.

3.1 Entry Model and Empirical Specification

In this setting, scientists observe all RFAs and must incur some costs to enter.[25] These costs are a function of, among other things, the scientific similarity between the individual's prior work and the RFA objectives, s. All potential entrants $i = \{1, 2, ..., N\}$ observe each RFA k, which are characterized by the total amount of funds available (Purse_k), the expected number of competitors ($\hat{n}_{ik} = \{1, 2, ..., N-1\}$), a vector of observable characteristics ($\mathbf{X}_k$; e.g., year, NIH Institute), and all other characteristics valued by scientists but unobservable to the econometrician (ξ_{ik}). A "revenue" function W and entry cost function C are based on these variables[26], such that each individual's expected payoff from entry is given by:

$$V_{ik} = W(\text{Purse}_k\,,\,\hat{n}_{ik}\,,\,\mathbf{X}_k\,,\,s_{ik}\,,\,\xi_{ik}) - C(\mathbf{X}_k\,,\,s_{ik}\,,\,\xi_{ik}). \tag{2}$$

The elasticity of science with respect to purse size is $\varepsilon \equiv \frac{\partial s_{ik}}{\partial \text{Purse}_k} = \frac{\partial V_{ik}}{\partial \text{Purse}_k} / \frac{\partial V_{ik}}{\partial s_{ik}}$. Scientists trade off the "market size" (Purse_k) for redirections, holding fixed their competitive expectations. To identify the partial derivatives necessary, I assume that W is linear in its parameters yielding the main estimating equation

$$\Pr(\text{Entry}_{ik}) = F(\text{Purse}_k) + G(\hat{n}_{ik}) + D(s_{ik}) + \mathbf{X}_k\beta_x + (\xi_{ik} + \mu_{ik}) \tag{3}$$

where μ_{ik} are i.i.d. mean-zero error terms that capture random noise in scientists' decisions.[27] Because each individual is extremely small relative to the full set of potential entrants, I assume that scientists are atomistic in the sense that they are oblivious to any effect their particular decision has on the rest of the sample.[28]

Note that whereas $\mathbf{X}_k$, s_{ik}, and ξ_{ik} enter Equation 2 through both W and C, they enter Equation 3 once and in a separable manner. Therefore, I simply interpret the partial derivatives as encompassing the cumulative costs and benefits of the independent variables. The estimate of ε is given by $\frac{F'}{D'}$, provided the units of the variables are comparable.

[25] Scientists also have the outside options of applying to the default open competitions or not at all. I assume the value of this outside option is fixed conditional on the covariates, which includes time- and individual-fixed effects.

[26] The W function also describes the rules by which the purse is allocated amongst entrants.

[27] This linear model is an admittedly strong assumption about both the nature of competition as well as scientists' utility over the outcomes. However, because the goal of this exercise is to estimate how scientists trade off redirection costs for expected resources - and not to simulate counterfactuals - these assumptions provide a transparent method for estimating the elasticity for the marginal entrant while controlling for first-order competitive effects.

[28] That is, there are no general equilibrium effects from any single scientist's decision. This is very reasonable given there are roughly 140,000 potential entrants, and about 16,000 unique individuals apply to the NIH in my data annually.

3.2 Approach to Endogenous Competition

The difficulty in estimating Eq. 3 is that, instead of competitive expectations ($\hat{n}_{ik}$), only the realized number of entrants (n_{ik}) is observed. Now, if $\mathbb{E}[\xi_{ik} \,|\, \text{Purse}_k, \hat{n}_{ik}, s_{ik}, \mathbf{X}_k] \neq 0$, then each scientist's likelihood of entry will be positively correlated. This will create an endogeneity problem if Equation 3 is estimated with n_{ik} instead of $\hat{n}_{ik}$, where estimates of G will be biased upwards and then, potentially, F downward.[29]

To address this issue, I use the procedure for estimating static strategic interactions outlined by Bajari et al. (2010).[30] $\hat{n}_{ik}$ can be estimated empirically if a variable exists that influences each individual's strategic choice *directly* but only influences others' choices via the *indirect* effect of those strategic choices. Estimates of $\hat{n}_{ik}$ can be obtained using variables that satisfy this "strategic exclusion restriction". In the model outlined above, similarity provides a valid instrument under the assumption that each scientist's similarity to an RFA is exogenous and does not directly influence any other scientists' behaviors.[31]

The underlying information structure assumed in this approach is that scientists know (1) the revenue and cost functions ($W(\cdot)$ and $C(\cdot)$), (2) the features of the RFA (Purse_k, $\mathbf{X}_k$), and (3) the number of other scientists (N) and the distribution of their similarities (s_{ik}) and unobservables (ξ_{ik}). Given this information, they can integrate over the distribution of predicted entry probabilities to form their expectations and then make their entry decisions. The estimation procedure is a similar process as follows:

1. Regress entry decision ($\text{Entry}_{ik} = \{0, 1\}$) on the vector of exogenous RFA characteristics: Purse_k , $\mathbf{X}_k$, s_{ik}
2. Predict entry probabilities, $\widehat{\Pr(\text{Entry}_{ik})}$ with parameters from Step 1
3. Sum predicted entry probabilities over each RFA, minus each individual's entry probability to estimate $\hat{n}_{ik}$, given by $\widetilde{n_{ik}} = \sum_{i' \neq i}^{N} \widehat{\Pr(\text{Entry}_{i'k})}$
4. Estimate Eq. 3 using $\widetilde{n_{ik}}$ in place of $\hat{n}_{ik}$

[29] This will depend on how correlated scientists preferences are for the unobservable features of RFAs (ξ_{ik}). As this correlation increases, so to will the correlation in entry probabilities across the sample, giving rise to a positive correlation between $\Pr(\text{Entry}_{ik})$ and n_{ik}.

[30] While there may certainly be dynamics with respect to each scientist's decision to purse a particular RFA (e.g., how would moving to topic A effect future research prospects?), the limited recurrence of RFAs and the massive scale of the default open competitions, which present a future option for funding, suggest that *competitive* dynamics are likely not first-order concern. Not at any time in discussions with scientists who have competed in RFAs did the notion of dynamic strategic interactions arise.

[31] One mechanism that may invalidate this assumption is if a scientist's likelihood of communicating with potential entrants is correlated with their similarity to an RFA. Anecdotal discussions with NIH applications who have competed in RFAs did not suggest this is relevant.

I estimate the regressions in Steps 1 and 4 via OLS, and assume that F, G and D are linear functions[32]. I will also include a scientist-fixed effect to control for stable differences across all scientists.

3.3 Data Construction and Summary Statistics

Estimating this model requires a dataset comprised of scientist-RFA pairs containing (1) all potential entrants, (2) RFA-level data on purse size and research objectives, and (3) each potential entrant's scientific similarity to the research objectives, which I will denote as *pmra*(Pubs.-RFA). To arrive at a close approximation to the full set of potential entrants, I include any individual that applied to the NIH from 2002 to 2015 totaling to 142,745 scientists. I then match each scientist (and their publication history) with each of the 394 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.[33]

In this sample, the average RFA has a first-year total purse size of \$2.89M (s.d. = 2.65) for both direct and indirect costs, attracts 38.2 entrants (s.d. = 42.3) and awards a total of 7.6 (s.d. = 7.3) grants with first-year direct costs totaling \$2.32M (s.d. = 2.65). For each of scientist-RFA pair, *pmra* was then used to calculate the scientific similarity between the scientist's prior publications and the research objectives of the RFA.

This research design resembles studies of firm entry where instruments, such as geographic distance, generate exogenous variation (e.g., Krasnokutskaya and Seim 2011). However, one difference between these studies and the approach of this paper is that when considering geographical distances, a firm must physically traverse the entirety of the distance. Here, although the *pmra* algorithm provides an estimate of this distance in this scientific sense, it is unclear whether scientists must actually travel the entirety of the distance since the bounds of an RFA are not explicitly defined. In the Appendix, 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. Still, the adjusted *pmra* is preferred because it more accurately captures the real changes to be expected by entrants, and permits a straightforward comparison to the effects identified in Section 4.

[32] i.e., $D(s_{ik}) = \beta_D \times \overline{s}_{ik}$, where $\overline{s}$ is the standardized value of s). In the Appendix results are shows where this is a log-linear function of the variables.

[33] Included in the vector of covariates $\mathbf{X_k}$ are indicator variables for the NIH Institute/Center that administered the RFA, the fiscal year the RFA expired, whether the RFA specifically requested "Collaborative" research be proposed, and whether the RFA solicited other non-R01 types of grants as well.

Table 1: Entry Model First-Stage

	(1)	(2)	(3)	(4)	(5)	(6)
Scientific	3.06	3.15	3.27	5.09	5.58	6.22
Similarity	(0.121)	(0.126)	(0.126)	(0.230)	(0.240)	(0.245)
RFA controls		Y			Y	
RFA F.E.			Y			Y
Scientist F.E.				Y	Y	Y

Note : Independent variable is a dummy indicating entry. All dependent variables are standardized, with all coefficients and standard errors scaled by 10^{-4}. mean(Entry$_{ij}$) = 2.68×10^{-4}. $N_{scientist-RFA} = 55,099,570$. N_{RFA} 394. RFA controls are a vector of characteristics including dummies for the fiscal year, the NIH Institute administering the RFA. Standard errors clustered within RFAs.

3.4 Results

First, for the estimation approach described above to be valid, it must be that the scientific similarity measure is a significant predictor of entry, and orthogonal to any features of each RFA that might be valued by scientists. Table 1 reports the first stage results regarding these two assumptions. As initially highlighted in Figure 1, the similarity score is highly predictive of entry, and there is not substantive changes to the magnitude of this relationship when including RFA-level controls. There is a relatively large increase in the magnitude of the coefficient when individual-fixed effects are introduced, indicating significant differences across scientists with respect to both their underlying propensity to enter an RFA and their average similarity to all NIH RFAs.

Table 2 presents the main results from estimating variants of Eq. 3 without competition controls (Cols. 1-2), with realized competition (Col. 3), and with the instrumented competition controls (Cols. 4-7). Without conditioning on competition, the results imply elasticities between 0.10 and 0.18. These will be underestimates to the extent that scientists are dissuaded by competition because the model does not capture the equilibrium effect whereby larger purse sizes will incentivize entry by competitors. In Column 3, where the naïve control of realized entrants is included (n_{ik}), the purse size of the RFA is no longer a significant predictor of entry and scientists appear to positively value increased competition. While not a de facto impossible observation, perhaps in a setting with very large externalities, the aforementioned identification issues suggest a cautionary interpretation. If this was in fact the nature of competition in this setting, it would persist when utilizing the instrumented competition measure. However, Columns 4-6 all portray a more familiar negative competitive effect.

Across the specifications where competition is instrumented, the coefficients indicate an elasticity between 0.19 and 0.36, with the preferred specification that includes scientist-

Table 2: Entry Model & Implied Elasticities

	(1)	(2)	(3)	(4)	(5)	(6)
Scientific Similarity	3.03 (0.117)	5.61 (0.234)	3.03 (0.118)	3.26 (0.125)	6.22 (0.246)	6.22 (0.246)
Purse	1.30 (0.460)	1.40 (0.415)	0.0213 (0.0386)	2.69 (0.556)	2.89 (0.502)	2.84 (0.386)
Realized Competition			2.88 (0.0489)			
Instrumented Competition				-1.63 (0.220)	-2.29 (0.175)	-2.37 (0.217)
Scientist F.E.		Y	Y		Y	Y
RFA controls		Y	Y			Y
ε	0.184 [0.05-0.34]	0.107 [0.04-0.19]	-	0.355 [0.20-0.54]	0.200 [0.12-0.29]	0.196 [0.13-0.27]

Note : Independent variable is a dummy indicating entry. All dependent variables are standardized, and all coefficients and standard errors are scaled by 10^{-4}. mean(Entry$_{ij}$) = 2.68×10^{-4}. $N_{scientist-RFA} = 55,099,570$. N_{RFA} 394.Standard errors clustered within RFAs. ε is the elasticity of science at the sample means/s.d.s per the coefficients on Scientific Similarity and Purse, with bounds in brackets using the 95% C.I. of coefficient estimates.

level fixed effects and RFA controls identifying an elasticity of 0.196. Not accounting for competitive expectations would have underestimated the elasticity by roughly 50%. At the sample means, an elasticity of 0.2 implies that scientists are indifferent between being 1 standard deviation more similar to the objectives of the RFA and a $5.72M increase in the purse size.

To provide some comparison to the estimates I generate below in Section 4, it is useful to transform this magnitude into expected grant dollars. First, I estimate the average win probability for marginal entrants by simply regressing a dummy variable that equals 1 if the scientist wins on a dummy variable that equals 1 if the scientist enters the RFA, including RFA-fixed effects. This regression identifies a coefficient of 0.199 (s.e. = 0.007), which when multiplied by $5.72M implies that scientists are indifferent between a project that is 1 s.d. more similar to their prior work and $1.14M in expected grant dollars. At the sample means, such a project would total $1.4 million annually, roughly four times the average grant size, which would make it an extreme outlier. Less than 0.5% of awards in the sample are this large.

3.5 Heterogeneity, Robustness and Extension

Understanding the heterogeneity in this elasticity is essential for understanding why certain individuals will be more or less likely to respond to both endogenous scientific progress and exogenous policy levers.

Table 3: Heterogenous Elasticities

	(1)	(2)	(3)	(4)	(5)
Panel A:					
Similarity Quintile	1 (least)	2	3	4	5 (most)
ε	2.254	0.641	0.085	0.035	0.020
	[n/a]	[0.101-10.5]	[0.028-0.204]	[0.018-0.061]	[0.016-0.025]
$N_{scientist-RFA}$	14,235,537	7,760,802	11,015,939	11,033,012	11,033,461
Panel B:					
Avg. Year of 1^{st} Pub.	1974	1988	1995	2000	2004
ε	0.110	0.123	0.162	0.260	0.295
	[0.082-0.143]	[0.092-0.160]	[0.116-0.217]	[0.151-0.399]	[0.147-0.514]
$N_{scientist-RFA}$	9,572,028	9,843,386	10,550,538	9,333,094	7,821,904

Note : Estimated elasticities by quintile of Scientist-RFA similarity (Panel A) or year of first publication (Panel B). Bounds in brackets are based on 95% C.I.s from the models ("n/a" if coefficients not significant at 95% C.I.), with the full table of results in the Appendix.

Nonlinear Elasticity—: The raw data presented in Figure 1 suggested a nonlinear relationship between similarity and a scientist's likelihood of entering an RFA. The figure shows that marginal increases in similarity seem to imply larger increases in the likelihood of entry *for scientists who are relatively more similar to begin with.* Interestingly, this is suggestive of concave costs. Table 3 Panel A presents five elasticity estimates, subsetting scientists based on their quintile of similarity scores.[34] The implied elasticities is clearly monotonically decreasing across the similarity quintiles - scientists closer to an RFA are more inelastic. The sample average estimate from the previous section (0.2) appears to be the mean of a very dispersed distribution that ranges two orders of magnitude across the similarity distribution (2.3-0.02). This is driven by the fact that, as the similarity between a scientist and RFA grows, similarity becomes much more important to scientists relative to the value of additional funds. This result is counter to most settings where resources eventually become scarce, and thus the costs of adjustment increases with the degree of those changes. This does not appear to be the case. In terms of policy implications, this result suggests that, to the extent these extra-marginal scientists can be selectively targeted, they may be more amenable to adjusting their direction.

Experience—: Life-cycle effects on innovation have received much attention from economists, particularly in the biomedical arena. However, most research to date has focused on purely vertical production effects (i.e., the *rate* of science, Levin and Stephan 1991) or how the increasing burden of fundamental knowledge changes educational investments, and therefore age of peak output over time (Jones 2010). Because I cannot directly observe scientists'

[34]Importantly, the dependent variables are standardized within each model to ensure that I compare relative changes within these subsamples.

ages, I investigate a highly correlated variable: years since first publication. This variable incorporates both classic life-cycle effects (e.g., investing less in human capital later in life) and other experience-related outcomes and incentives (e.g., resource acquisition, learning). This is a very policy-relevant parameter because, following the logic of Jones' knowledge burden hypothesis (2010), pure aging effects evolve over time based largely on growth in the fundamental knowledge base, something much less amenable to intervention. Shaping policies based on a scientist's experience, however, takes this as a given.

Table 3 Panel B displays the estimated elasticities across the experience distribution. There is a clear pattern of more experienced scientists being less elastic with respect to these funding opportunities. The most experienced scientists in this sample, who's mean year of first publication is 1974 are one third as elastic ($\varepsilon = 0.1$) as the youngest scientists ($\varepsilon = 0.3$) who first publish around 2004.[35]

The implications of this pattern are certainly context-dependent. Consider the case where experience and ability are positively correlated. Here, a manager intent on directing production *in a particular direction* may need larger awards to incentivize higher ability scientists. Or, in the case where scientists must search for new questions to pursue and the returns to a given scientist-question match is decreasing in experience - it is better for early-stage researchers to find their "ideal" question given the longer remaining time for them to pursue it - then it is in the manager's favor that it is cheaper to incentivize redirections amongst these younger individuals.

Specification Robustness—: Robustness tests in the Appendix include results using the raw *pmra* score (sans the adjustment described in Section 3.3), only publications that occur within 5 years of the focal RFA, controls for publication counts, alternative transformations of *pmra* scores (average, median) and a log-linear formulation of Equation 3 (i.e., $F(\text{Purse}_j) = \beta \times \log(\text{Purse}_j)$). Across the difference specifications the main results persist, and in each case, not controlling for competitive effects leads to significantly smaller elasticity estimates.

Identification Robustness—: In order to explore the main assumption $\mathbb{E}[\xi_{ik} \,|\, \text{Purse}_k, \hat{n}_{ik}, s_{ik}, \mathbf{X}_k] = 0$, the Appendix details two tests. In the first, a jackknife approach is taken where the entry model is re-estimated 394 times, leaving out a single RFA one at a time. The distribution of elasticities identified is very narrow (mean = 0.20, s.d. = 0.004). Furthermore, there is no meaningful correlation between each jackknife estimate of the elasticity and the purse size of the RFA removed. If purse sizes were in fact correlated with unobservables, I have expected to systematically under- or overestimate the elasticity

[35]These analyses only include scientists with at least one publication prior to the expiration of the RFA, again using the maximum *pmra* score generated.

when larger RFAs were excluded from the sample.[36] Furthermore, the main results are robust to excluding all RFAs from the top 5% of the purse distribution, indicating that these particular competitions are not driving the results.[37]

In the second test, I examine the value of competition implied by the model. The argument is based on the fact that, if $\mathbb{E}[\xi_{ik} \mid \text{Purse}_k, \hat{n}_{ik}, s_{ik}, \mathbf{X}_k] = 0$, then the model should accurately identify the dollar value (in terms of purse size) that scientists place on the addition of a marginal competitor ($\partial \hat{n}_{ik} = 1$). Unlike the elasticity of science, this "elasticity of competition" can directly be approximated from the data and compared to the model's estimate. In both cases, it appears that scientists are indifferent between a roughly \$80,000 increase in purse size and the reduction of one expected competitor. Combined with the MeSH-level results from Section 2.3, this evidence indicates there is no meaningful correlation between the creation or features of RFAs and the scientific topics they target.

Adjustment Costs vs. Production Costs—: All research grants are implicitly funding two aspects, adjustment and production. That is to say, the costs of these grants scale in two dimensions: rate (e.g., number of new publications) and direction (e.g., the similarity of new publications). The elasticity estimate can be extended in order to understand what share of the total costs observed in this setting are due to either of these dimensions. In the Appendix, a simple thought experiment suggests that costs that scale with a new project's direction account for roughly 78% of total costs. In other words, changes in direction impact costs 4.5 times ($=1/(1-0.78)$) more so than changes production of the same relative size. Given the lack of information currently known about such costs, hopefully this estimate can spark new theoretical and empirical investigations.

3.6 Elasticity of Science in the Aggregate

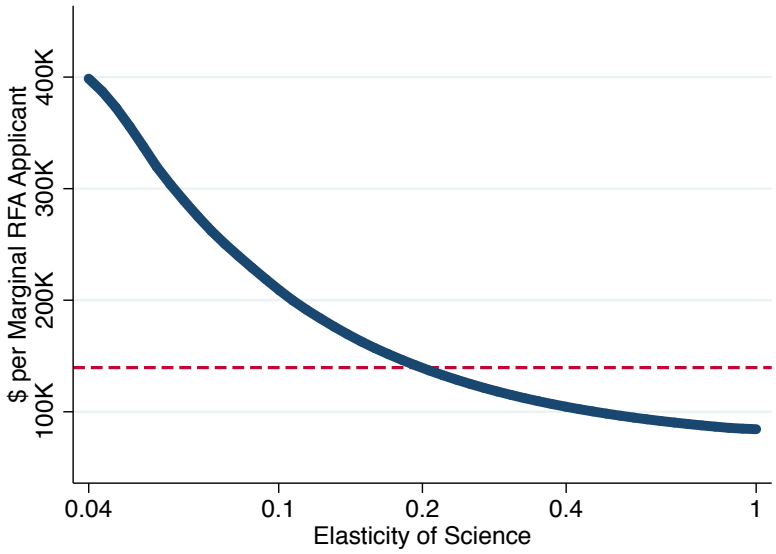
The results thus far have emphasized the costs of incentivizing redirections for each particular scientist in isolation. But given that the base of scientists is approximated here at 140,000 individuals, and these individuals appear sensitive to competitive effects, it is useful to understand how the magnitude of this parameter influences aggregate outcomes. A useful metric to explore is the marginal costs of incentivizing 1 additional applicant to an RFA.

As a frame of reference, the baseline model indicates that 1 additional scientist will enter an RFA (in expectation) if the purse size is increased by \$139,616 from the average. This

[36] Under the assumption that the entry function is truly linear in the Purse variable.

[37] Notably, variation in purse sizes amongst the smaller RFAs is much more likely to be due to exogenous events such as budget shocks as indicated in discussions with NIH staff.

Figure 5: Marginal Costs of Inducing Applications



Note : Costs in terms of RFA purse size; local polynomial smoothed. Dashed line indicates the implied marginal cost at the average elasticity identified ($\varepsilon \approx 0.2$), which is $139,616.

magnitude is roughly twice the size of a crude estimate of the expected marginal benefit of entry.[38] To explore how this cost varies with the elasticity of science, the entry model is re-estimated for a range of elasticity values. The underlying methodology of this exercise is outlined in the Appendix.

Figure 5 displays the implied costs per applicant for a range of elasticity estimates spanning +/- 5 times the baseline mean of roughly 0.2. This is within the range of estimates identified in the heterogeneity analyses earlier. In the polar cases, an elasticity of 0.04 suggests a cost per applicant of nearly $400,000 while an elasticity of 1 suggests costs of roughly $84,000 per applicant. On average, this implies that a 10% increase in the elasticity results in a 5% decrease in the marginal cost of incentivizing RFA applications.

However, this relationship does not appear to be a linear. As scientists become increasingly elastic, the costs of inducing entry begin to plateau. Intuitively, this plateau is near the implied expected value of entry, highlighting the strategic nature of scientists decisions: more elastic scientists are more willing to adjust their trajectories to pursue an RFA, but if all scientists are more elastic, *and all scientists are aware of this*, then their competitive expec-

[38]The marginal win probability is approximately 0.2. Given this is virtually identical to the average win probability, I assume that marginal winners receive the average award size to approximate the expected value of entry at $68,000.

tations will push them to not pursue these grant funds without net positive expectations.[39] This latter point is certainly policy relevant. It indicates that, should policies be pursued to promote increased movement across fields (increased elasticity), there is a limit to how much these policies can lessen the costs associated with pulling scientists into new fields.

4 Is there a Premium for Trying to Attract New Brains?

4.1 Motivating Framework

The goal of this section is to investigate whether the costs of incentivizing redirections has a first order effect on the size of grants awarded at the NIH - are these costs "large"? The logic is that if it is costly for scientists to adjust the direction of their work, and RFAs require larger adjustments relative to the default open grants, then the expected value of RFAs should be larger.[40] Figure 6 Panel (a) provides initial evidence that RFAs require individuals to adjust their trajectories relative to open grants; RFA applications are 25% less similar on average.[41]

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, discussions with NIH staff made clear that the division of funds takes into account the number and quality of applications.[42] Competition influences awards. And as evidenced in Panels (b-c) of 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

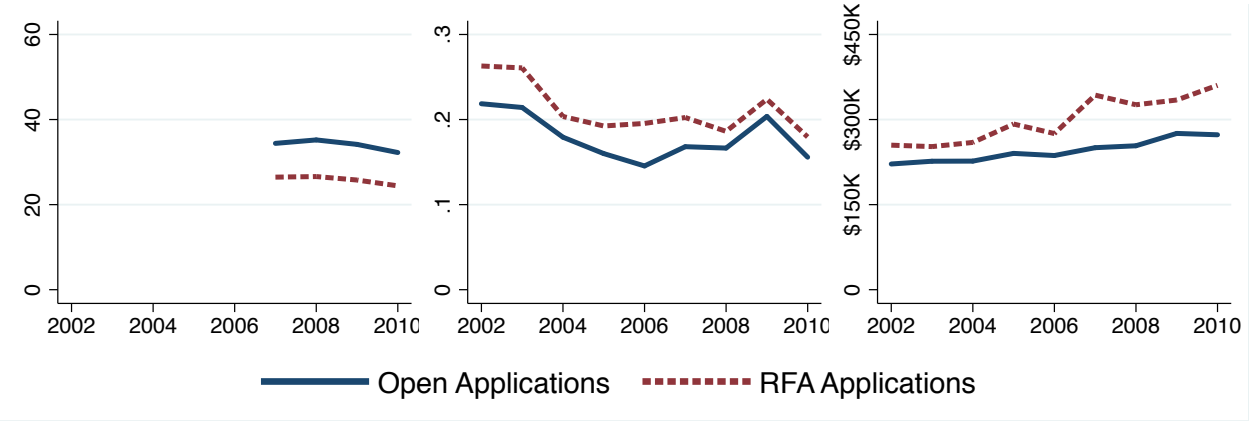
[39] The wedge between the observed plateau at about $80,000 and the implied expected value of entry at $68,000 sheds some light on the fixed costs of entering these competitions.

[40] The Appendix 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.

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

[42] 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.

Figure 6: Trends in Outcomes, RFA versus Open Applications



(a) Similarity **(b)** Win Probability **(c)** Award Size

Note : RFA-Open correlations are 0.97 for similarity, 0.90 for win probability and 0.92 for award size. Data availability restrict the set of applications for which similarity scores can be calculated.

are more willing to redirect, more likely to win, and win larger awards). Then, as outlined in the Appendix, comparing the relative size of the redirection effect to the premium provides another way of identifying the focal elasticity: $\frac{\text{Similarity}_{RFA} - \text{Similarity}_{Open}}{\text{Expected Value}_{RFA} - \text{Expected Value}_{Open}}$.

Notably, this research design relies only on variation between RFA and open grants. Thus, while it identifies the causal effect of RFAs on the redirection and payoff outcomes, I can only directly link these two results by assuming that no other differences between the mechanisms are relevant, which may not be the case.[43] This assumption is not testable in the data, but I still report these implied elasticities because, when contrasted to the estimates from the entry model, they can inform to what extent the premium can plausibly be explained by the inelasticity of science.

4.2 Research Design

As a first pass, I could regress the three outcomes $y = \{$scientific similarity, win probability, award size$\}$ for scientist i's application j on an indicator for whether the application was

[43]Besides any redirection effect, these two funding channels differ in the nature of competition, which may play into scientists' preferences. 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 this option is much easier for open grants compared to RFA grants.

submitted to an RFA ($\mathbf{1}\{\text{RFA}_{ij}\}$) as follows:

$$y_{ij} = \alpha + \beta \mathbf{1}\{\text{RFA}_{ij}\} + \epsilon_{ij}. \tag{4}$$

From these regressions, the β coefficient can be used to describe the RFA redirection effect (change in similarity) and premium (change in expected value).[44] But this simple difference suffers from two drawbacks: (1) it is unclear that the average open application is the appropriate counterfactual for RFA applications, as is implied by Eq. 4, and (2) 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 the Appendix 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.

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 also ensures that applications of equivalent scope/scale are compared to one another.[45]

Rewriting Eq. 4 to include the fixed effects for individuals (i) and time-variant research areas (st) and the covariates ($\mathbf{X_{ijst}}$), the main estimating equation is:

$$y_{ijst} = \alpha_i + \beta \mathbf{1}\{\text{RFA}_{ijst}\} + \boldsymbol{\gamma}\mathbf{X_{ijst}} + \tau_{st} + \epsilon_{ijst}. \tag{5}$$

[44] The redirection effect will be directly estimated as $\beta_{\text{similarity}}$. 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 $\frac{(\bar{w}+\beta_{\text{win}})\times(\bar{a}+\beta_{\text{award}})}{\bar{w}\times\bar{a}}$.

[45] 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.

Equation 5 is estimated using OLS for the three outcomes: log-transformed similarity scores, a binary win indicator, and log-transformed first-year direct costs.[46]

4.3 Summary Statistics & Results

For the focal sample of applications included in this analysis, the average application is successful 23% of time, with an average first-year award size of $298K. This sample includes 12,852 unique scientists, 39,741 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 do in fact induce scientists to adjust the trajectory of their research, Table 4 estimates the similarity differences between the open and RFA competitions following Equation 5, introducing more rigorous controls across the columns. 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 individual's 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. The models with scientist-fixed effects suggest this difference is roughly 10-15%.

Table 4 also presents estimates of the win probability and award size differences. Across all specifications RFA applications are more likely to be awarded funds, and more likely to receive a larger amount of funds. The magnitudes indicate an RFA premium of roughly 65% relative to the expected payoff of open applications. This implies that scientists are indifferent between approximately $270,000 awarded via an open grant, and $445,000 awarded via an RFA. That this implied estimate is larger than the observed average RFA grant size in this sample ($380,000) suggests that some positive selection is occurring, emphasizing the limitation of comparing the unadjusted statistics in this setting.

Proceeding with the aforementioned caveat, these estimates imply an elasticity of about 0.2, with the most preferred specification (Table 4, Col. 5), implying 0.16. This suggests scientists are indifferent between a 1 s.d. adjustment to more similar science and a $870,000 larger grant. These magnitudes are very similar to those from the previous section, which found a 1 s.d. redirection was worth roughly $1 million. In other words, the (in-)elasticity

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

Table 4: Differences in Outcomes for RFA versus Open Applications

	(1)	(2)	(3)	(4)	(5)
Panel A: y =log(Similarity)					
1{RFA}	-0.232	-0.238	-0.149	-0.110	-0.102
	(0.00875)	(0.00874)	(0.00878)	(0.0152)	(0.0179)
$N_{\text{Applications}}$	39741	39731	21495	36237	18582
$N_{\text{Scientists}}$	26723	26716	8480	24864	7408
mean(Similarity$_{\text{open}}$)	33.50	33.50	34.30	33.40	34.30
Panel B: y =Pr(Win)					
1{RFA}	0.111	0.0749	0.0706	0.121	0.0883
	(0.00650)	(0.00496)	(0.00876)	(0.00961)	(0.0182)
$N_{\text{Applications}}$	39756	39746	21502	36250	18588
$N_{\text{Scientists}}$	26734	26727	8483	24874	7411
mean(Pr(Win)$_{\text{open}}$)	0.22	0.22	0.21	0.22	0.21
Panel C: y =log(Award \$)					
1{RFA}	0.240	0.199	0.197	0.0587	0.167
	(0.0149)	(0.0112)	(0.0229)	(0.0235)	(0.0504)
$N_{\text{Applications}}$	9743	9738	1944	8447	1182
$N_{\text{Scientists}}$	8698	8693	899	7605	555
mean(Award \$$_{\text{open}}$)	270646	270670	288053	271693	284237
Application Controls		Y	Y	Y	Y
Scientist F.E.			Y		Y
Research Area-Time F.E.				Y	Y

Note : App. controls include peer review score and funds requested. Standard errors clustered within scientists.

of science identified by the entry model can account for virtually all of the difference in expected payoffs between RFA and open grants.

5 Do the New Brains Produce as Intended?

The results of Sections 3 and 4 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 scientists' work.[47] This section uses the rank-order funding process of RFAs to identify the impact of grant receipt on the rate and direction of scientists' publications.[48]

5.1 Data & 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. I identify publications where the scientists was a primary investigator by

[47]The key issue is that 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.

[48]The following results are conservative in that they only compare outcomes conditional on application. The lack of a clear counterfactual for all appliers prevents a robust analysis.

proxying for this role based on whether the 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 the head of the laboratory or chiefly responsible for the design of the study.

Publication Direction—: To assess the direction of each scientist's publications, I use the *pmra* algorithm to score the similarity between each applicant's publications and the research objectives of the RFA applied to. I then take the logged difference of the average similarity scores of all publications before and after the award decision to generate the focal measure of direction.[49] Taking the average eliminates any variation due to changes in publication rates, and by logged difference the estimated coefficients can be interpreted approximately as percent changes.

5.2 Regression Discontinuity Design

The RFA award process is as follows: first, applications receive quality scores from peer review panels consisting of scientists from relevant fields; second, applications are sorted based on these review scores; finally, funds are awarded in imperfect rank-order at the discretion of NIH staff until the RFA's budget constraint binds. The approach of this empirical analyses will be to compare the outcomes for marginally funded and non-funded applications, under the assumption that the budget constraint is exogenous, while also accounting for the imperfection in the rank-order funding.[50]

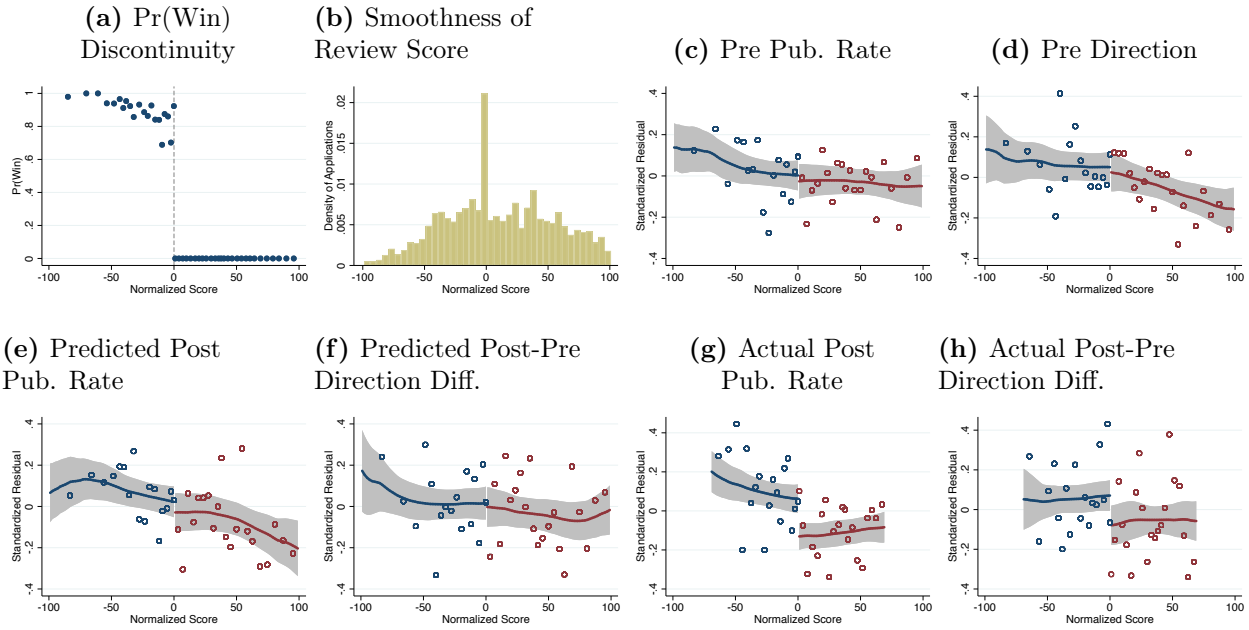
I follow the approach of Jacob and Lefgren (2011), who study open grants and their effect on publication output. I estimate the local average treatment effect (LATE) of NIH grant receipt around the *funding threshold*, defined as the worst-ranked successful application. Because not all applications on the "successful" side of the threshold receive grant awards - treatment compliance is 85% - the LATE is estimated using the so-called fuzzy regression discontinuity design. Following Jacob and Lefgren (2011), I generate a normalized score by centering applications within each RFA around the review score of the threshold application. Because lower review scores indicate higher quality, the normalized review score z for application j to RFA k is defined as $z_{jk} = \text{score}_{jk} - \overline{\text{score}}_k$, where $\overline{\text{score}}_k$ is the review score of the RFA k's threshold application.

I estimate variants of Eqs. 6, where a binary indicator of grant receipt, Win_{jk}, and the publication outcomes Y_{kj} are regressed on flexible functions of the normalized score and

[49] This focal measure is: $\log(\frac{mean(pmra_{(\text{Post Pub}_p,\ \text{RFA Obj.})})}{mean(pmra_{(\text{Pre Pub}_p,\ \text{RFA Obj.})})})$, where Pre Pub_p and Post Pub_p index the applicant's set of publications before and after the RFA award decision, respectively.

[50] The Appendix details the NIH award policies that make this research design possible.

Figure 7: Discontinuity Diagnostics and Raw Outcomes



Note : Rate and Direction outcomes, and prediction methods described in-text. The spike at 0 in Panel (b) is due to the fact that all RFAs by construction have at least one application with a score of 0. For Panels (c-h), y variables are displayed as standardized residuals conditioning on RFA-fixed effects; lines are local linear estimates (95% C.I. shaded), with equally binned circles at means.

RFA-fixed effects. The instrumental variable in this setup is a dummy variable that equals 1 if the application is below the funding threshold, denoted by $\mathbf{1}\{z_{jk} \leq 0\}$.

$$\text{Win}_{jk} = \alpha + \beta \mathbf{1}\{z_{jk} \leq 0\} + f(\text{z}_{jk}) + \sigma_k + \mu_{jk} \tag{6a}$$

$$Y_{jk} = a + \gamma \text{Win}_{jk} + g(\text{z}_{jk}) + \varsigma_k + \nu_{jk} \tag{6b}$$

Figure 7 Panel (a) plots this discontinuity in award probability across the distribution of z values. Note that award probability is zero to the right of the threshold (worse applications) by construction. I follow Calonico, Cattaneo and Titiunik (2014) to jointly estimate Equations 6 as fuzzy regression discontinuity (RD) where f and g are local linear functions independently estimated on either side of the threshold, and coefficient and confidence interval estimates are adjusted to account for bias that arises under different bandwidth specifications.

The two key assumptions of this research design are that z_{jk} cannot be manipulated to move an application to/from either side of the threshold for reasons correlated with Y_{jk}, and the

location of the threshold within each RFA is plausibly exogenous to changes in Y_{jk} at that point of the z distribution. The first assumption is well-founded in this setting. Applicants themselves can certainly not manipulate their review scores, because all peer review and funding decisions are made without their input. And the peer reviewers assign scores prior to any funding decisions. It is likely that they have informed expectations about where the funding threshold will lie prior to assigning scores. However, funding outcomes are *relative to the realized score rankings*, in that grants are awarded to the top applications based on the budget constraint. Applications above some score are not inherently funded. Thus, there is no way to manipulate an application's score without then altering the likelihood that another application win - the incentive is always to rate higher quality applications better.[51]

The most pertinent concern here is the location of the threshold. The primary determinant of this threshold is the RFA's budget constraint, which is assumed to be orthogonal to any relevant feature of the RFA.[52] However, discussions with staff and the observation of imperfect rank-order funding indicate that it is possible for the officials overseeing an RFA to influence the location of the threshold. This can be done by skipping applications, withholding funds, or lobbying for a larger budget. This may result in an ordering of applications where the marginally funded and unfunded applications are not observationally equivalent conditional on the control variables; the threshold application may not be "random".

To investigate the extent to which this, or any other feature, may contaminate the analyses, Figure 7 Panels (b-f) present a range of standard tests for manipulation. Panel b plots the density of the normalized review score around the threshold. If thresholds are systematically placed at portions of the score distribution characterized by large jumps in unobservable quality, we would expect to see discrete changes in this density. No first-order changes are evident, noting that the spike at zero is because all RFAs have at least one application with this value by construction.[53]

Panels (c) and (d) plot the pre-period versions of the dependent variables. For publication counts, this is simply the total sum of first- or last-author publications prior to the award decision. For the direction measure, this is the average similarity of all pre-period publications relative to the research objectives of the RFA applied to. Panels (e) and (f) plot the predicted values of the outcomes of interest using a large vector of application and scientist observables and a LASSO estimator. In all cases, scientists and their applications look very

[51]Furthermore, Li (2017) shows evidence that gains due reviewer's expertise tend to dominate any losses associated with biases from a preference for work similar to one's own.

[52]This assumption is motivated by the prior results where initial RFA outlays do not appear to be meaningfully correlated with any aspect of RFAs that scientists' would value.

[53]This also prevents formal statistical testing due to the discrete change in the distribution.

similar with respect to these variables on both sides of the funding threshold.

To preserve sample size for the estimation, the bandwidth for the main specifications is set at 100, which is the smallest range that includes the border applications from each RFA in the sample. In the Appendix, I contract and expand this bandwidth to explore how this choice affects the results.[54]

To summarize, Figure 7 Panels (a-f) illustrate the sharp discontinuity in funding probability and suggest that applications near this threshold appear to be very similar on observable characteristics. As a further test of the research design's validity, I will present so-called "donut" results where the applications directly bordering the threshold are dropped from the analyses as they are likely the applications most at-risk of manipulation.

5.3 Summary Statistics & Results

I evaluate outcomes for scientists that applied to an RFA between 2001 and 2009, including only RFAs where there is at least one winning application, not all applications win, and only examine outcomes for scientists' whose application received a review score, as this is required to generate the running variable. The final sample includes a set of 141 RFAs and roughly 2,600 applications. The Appendix includes a table with summary statistics for this sample.

Figure 7 Panels (g-h) plot the realized outcomes across the funding threshold and show significant discontinuities in both the rate and direction outcomes. These figures indicate that applications marginally across the threshold appear to have more publications in the post period, and have larger relative changes in similarity compared to the objectives of the focal RFA. Table 5 presents these results in regression format. For comparison, the Table also includes simple OLS specifications with a global, linear control for the application's review score. Panel (a) includes all applications, and Panel (b) drops the border applications that are either the last-funded or first-unfunded.

In all cases, the fuzzy RD model estimates positive and significant LATEs for the rate and direction outcomes. I find that exogenous grant receipt around the threshold leads applicants to publish roughly 2.5-3.5 additional first/last author publications in the following years (40% of the mean). The estimates also indicate that, relative to marginal losers, marginally successful applicants change the direction of their work to be roughly 15-40% more similar to

[54]To get a sense of the quality range of 100 points in this metric, this corresponds to a 20% lower/higher review score based on the scoring criterion used by the NIH.

Table 5: Award Effects on Rate and Direction Outcomes

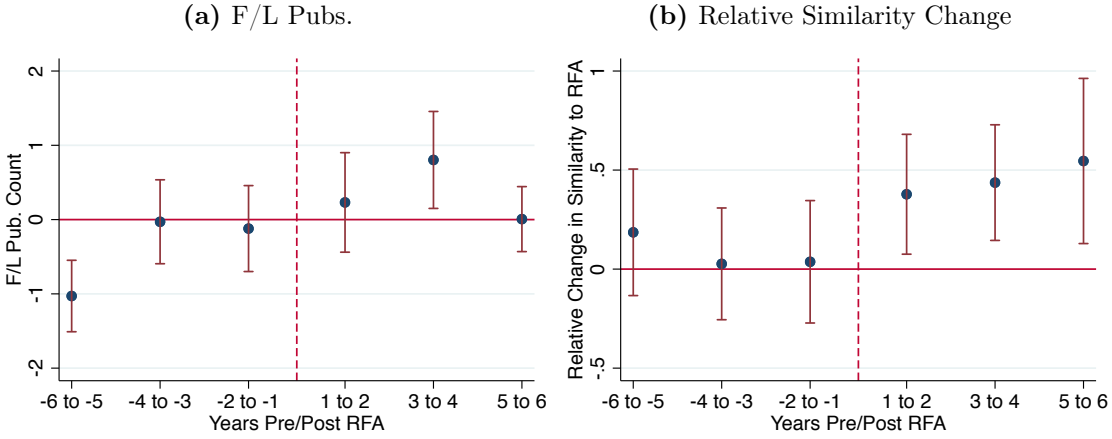
	(1)	(2)	(3)	(4)	(5)	(6)
	F/L Author Pub Count			Relative Similarity Diff.		
Panel A: Main Sample						
Win	1.076*	3.693*	2.941	0.0101	0.436	0.434
	(0.575)	(1.908)	(1.327)	(0.0302)	(0.102)	(0.0864)
N	2408	2408	2408	1822	1822	1822
mean Dep. Var.	7.058	7.058	7.058	0.219	0.219	0.219
Panel B: Excl. Border Apps.						
Win	1.425	3.352	2.447	0.00950	0.180	0.155
	(0.644)	(1.177)	(0.814)	(0.0334)	(0.0592)	(0.0501)
N	2129	2129	2129	1628	1628	1628
mean Dep. Var.	7.076	7.076	7.076	0.246	0.246	0.246
RFA F.E.	Y		Y	Y		Y
Spec.	OLS	RD	RD	OLS	RD	RD

Note : Dependent variables and specifications described in text. Standard errors clustered within RFAs.

the objectives of the focal RFA. These magnitudes are of the same order as the RFA versus open difference estimated in Section 4, which ranged from 10 to 20%.

Panel (b) of Table 5 indicates that the main results are not particularly sensitive to the exclusion of applications immediately bordering the threshold. One exception is that the LATE estimate for the redirection effect is significantly different when the border applications are excluded. This suggests that either these marginal applications are particularly sensitive to grant receipt, or that NIH staff are influencing the threshold location in a manner that favors funding towards scientists that would undertake significant redirections with or without grant funds. Still, a significant effect persists when these applications are excluded.

Figure 8: Fuzzy Regression Discontinuity Results, Event Studies



Note: Plots the LATE estimates (1 model per point, with 95% C.I.) from Eqs. 6.

Figure 8 presents results from event study versions of these analyses. In Panel (a), the outcome is the count of First/Last Author publications. In Panel (b), the similarity difference

is based on publications prior to 6 years before the RFA (the earliest relative time period examined) and each successive year's publications.[55] The results show that prior to award decisions, there are no major differences between marginal winners and losers, except that marginal winners appear to have slightly lower publication rates in one of the pre-periods. The event study indicates that the LATE identified in the main results is driven by new publications in the 3 to 4 years post award decision, which is in line with the average four year lifespan of these grants. With respect to changes in the direction, scientists' post-award work becomes immediately and increasingly aligned with the RFA's objectives over time, suggesting that the redirections persist, at least in the short run.[56]

Overall, the estimates show that the receipt of a grant is a key determinant of whether or not scientists pursue the topics targeted by the RFA. Not only does this provide evidence that NIH grants can stimulate new research on specific topics, it indicates that scientists responding to RFAs are not simply seeking resources for activities that will be undertaken regardless of the application's outcome. This is further evidence that, if not for RFAs and the funds they provide, scientists would not pursue these new directions.

6 Discussion

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, pp. 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 one of the most skilled sets of the economy, academic scientists. My results indicate that, at least in this particular market, these costs are likely of first order concern. Furthermore, the role of these costs will differ whether one

[55] This metric is defined for each period $y = \{(-6,-5),(-4,-3),(-2,-1),(1,2),(3,4),(5,6)\}$ by: log(mean($pmra_($Publications$_{py}$, RFA))) - log(mean($pmra_($Publications$_{pre-6}$, RFA))), where Publications$_{py}$ indexes the applicants publications in each period, and Publications$_{pre-6}$ includes all publications prior to 6 years before the RFA.

[56] The Appendix also illustrates how these effects vary across the distribution of publication quality and magnitudes of similarity changes. On average, winners' new publications are of roughly comparable quality to those generated by marginal losers. For the direction outcome, marginal winners are more likely to undertake the largest relative redirections observed in sample. Results are also presented where the score bandwidth is varied. In all cases, smaller bandwidths identify larger, but in general not statistically different, estimates.

is looking to induce *a single scientist* to change directions in isolation (as in Section 3.4) or instead understand the aggregate costs of getting *any scientist* to pursue a certain direction (as in Section 3.6).

Understanding this elasticity is essential for policymakers and managers alike, because the vast majority of scientists based at public and non-profit institutions choose their own pursuits with minimal oversight. This system has arisen for good reason: uninformed funders are willing to relinquish control in order to leverage scientists' private information (Aghion, Dewatripont and Stein 2008). However, it is not clear a priori that the allocation of funds thus decided by scientists themselves will be optimal from society's perspective (Dasgupta and David 1994).[57] To resolve such concerns, Aghion, Dewatripont and Stein (2008) suggest the use of mechanisms analogous to the RFAs I study here.

An obvious next question is how the presence and magnitude of these costs influences the optimal allocation of research funds. A few theoretical models have been developed to formally study this question in the biomedical realm specifically (e.g., Lichtenberg 2001). However, none of these models reckon with the type of costs emphasized here. Future work on this topic should be worthwhile. Similarity, the overall welfare implications of these results are not immediately obvious. For instance, although it may appear that the inelasticity I identify could prevent scientists from efficiently choosing their direction from society's view, they may in fact be a source of diversity that counterbalances other distortions arising from market dynamics (Acemoglu 2011) or racing incentives (Stephan 1996).

Specifically in the context of funding mechanisms for basic research, these results are relevant to the ongoing "people versus projects" debate regarding the optimal structure of grants.[58] 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 as to the specific use of funds. A frequent argument in support of these people-centered arrangements is that they provide incentives to be "creative". The apparently large costs of redirections I identify provide another, straightforward channel as to the value of people-centered funding: because project-centered mechanisms must cover both adjustments and the actual production, they limit the extent to which scientists can adjust their work *even when they are aware of a new, optimal direction.*

[57] 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, pp. 506).

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

References

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

Acemoglu, Daron. 2011. "Diversity and technological progress." In *The Rate and Direction of Inventive Activity Revisited*. 319–356. University of Chicago Press.

Aghion, Philippe, Antoine Dechezleprêtre, David Hemous, 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 Stein. 2008. "Academic freedom, private-sector focus, and the process of innovation." *The RAND Journal of Economics*, 39(3): 617–635.

Arrow, Kenneth. 1962. "Economic welfare and the allocation of resources for invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors*. 609–626. Princeton University Press.

Azoulay, Pierre, Christian Fons-Rosen, and Joshua S Graff Zivin. 2016. "Does science advance one funeral at a time?" National Bureau of Economic Research.

Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso. 2011. "Incentives and creativity: evidence from the academic life sciences." *The RAND Journal of Economics*, 42(3): 527–554.

Bajari, Patrick, Han Hong, John Krainer, and Denis Nekipelov. 2010. "Estimating static models of strategic interactions." *Journal of Business & Economic Statistics*, 28(4): 469–482.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica*, 82(6): 2295–2326.

Crane, Diana. 1969. "Social structure in a group of scientists: A test of the" invisible college" hypothesis." *American sociological review*, 335–352.

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." Vol. 33, 93–129, Elsevier.

Jacob, Brian, and Lars Lefgren. 2011. "The impact of research grant funding on scientific productivity." *Journal of Public Economics*, 95(9): 1168–1177.

Jones, Benjamin F. 2010. "Age and great invention." *The Review of Economics and Statistics*, 92(1): 1–14.

Krasnokutskaya, Elena, and Katja Seim. 2011. "Bid preference programs and participation in highway procurement auctions." *The American Economic Review*, 101(6): 2653–2686.

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

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.

Levin, Sharon, and Paula Stephan. 1991. "Research productivity over the life cycle: Evidence for academic scientists." *The American Economic Review*, 114–132.

Lichtenberg, Frank. 2001. "The allocation of publicly funded biomedical research." In *Medical care output and productivity.* 565–590. University of Chicago Press.

Li, Danielle. 2017. "Expertise versus Bias in Evaluation: Evidence from the NIH." *American Economic Journal: Applied Economics*, 9(2): 60–92.

Lin, Jimmy, and W John Wilbur. 2007. "PubMed related articles: a probabilistic topic-based model for content similarity." *BMC Bioinformatics*, 8(1): 1.

Manso, Gustavo. 2011. "Motivating innovation." *The Journal of Finance*, 66(5): 1823–1860.

Merton, Robert K. 1973. *The sociology of science: Theoretical and empirical investigations.* University of Chicago press.

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

Nelson, Richard R, et al. 1962. *The Rate and Direction of Inventive Activity: Economic and Social Factors.* Princeton University Press.

Oi, Walter Y. 1962. "Labor as a quasi-fixed factor." *Journal of Political Economy*, 70(6): 538–555.

Schmookler, Jacob. 1966. *Invention and Economic Growth.* Harvard University ress, Cambridge, MA.

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): 835–853.

Torvik, Vetle I, and Neil R Smalheiser. 2009. "Author name disambiguation in MEDLINE." *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 3(3): 11.

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

Appendix

Appendix A Additional NIH Details

A.1 NIH Funding Channels

The application process generates groups of applications in three dimensions, which following Azoulay et al. (2015b), I refer to as Disease (per the Institute the application is submitted to for funding, i.e., National Cancer Institute), Science (per the peer review group charged with evaluating the quality of the application, i.e., Bacterial Pathogenesis) and Time (per the fiscal year the application is submitted). Applications are peer reviewed per their Science-Time group and compete for funding per their Disease-Time group. These Disease-Science-Time (D-S-T) groupings are used to construct counterfactuals in the analysis of Section 4. Importantly, there are no explicit restrictions on the types of science that may be submitted to these competitions, so long as it fits within the NIH's broad objectives. "The NIH's mission is to uncover new knowledge that will lead to better health for everyone. Simply described, the goal of NIH research is to acquire new knowledge to help prevent, detect, diagnose, and treat disease and disability."[1]

Because RFAs are reviewed by a stand-alone peer review group convened just for the RFA, they are technically each unique D-S-T groupings. Thankfully, the NIH Center for Scientific Review (CSR) publishes broader sets of peer review groups, which RFA review groups are assigned to in the administrative data. These "Integrated Review Groups" allow me to match each RFA to a corresponding D-S-T, with the "S" now defined by the Integrated Review Groups. See the CSR for further details.[2] These D-S-Ts are used to define the time-specific research area fixed effects included in the analyses of Section 4.

A.2 RFA Generation and Empirical Relevance

Discussions with the NIH staff responsible for managing RFAs indicated two major forces: political and programmatic. As evidence of the political influence, the NIH's annual Congressional appropriation bills regularly include "soft earmarks," where members request research on specific topics. Hegde and Sampat (2015) find evidence that these diseases referenced by Congress often appear as the focus of grants awarded via RFAs.[3] Whether this relationship would lead my elasticity estimates to be over- or underestimates of the population average depends on whether Congressional preferences are correlated with positive or negative features of the science (from scientists perspective). Hegde and Sampat's (2015) findings indicate that disease advocacy lobbying is a key driver of Congressional preferences, and that this lobbying

[1] This is an approximation of the official funding process, which is outlined here: https://goo.gl/blLuuU, accessed July 12, 2017.

[2] https://goo.gl/PmYp2P, accessed July 12, 2017.

[3] This channel of influence was confirmed by NIH staff, who noted that these formal requests were often reinforced by direct communications (i.e., phone calls with staffers, on-site visits).

is positively correlated with both disease burden and scientific opportunities, which would suggest a positive correlation between these unobservables, scientists' preferences and RFA generation. However, it is important to note that although Congress may request research on, for example, a specific disease, the nature of RFA mechanisms and their historical use by NIH staff is such that they rarely target a single topic. In effect, this means that Congress's influence will be mixed and muted by the programmatic concerns detailed next. As example, consider the case of the Zika virus outbreak in 2016-17. Unsurprisingly, following this event, the Zika virus appeared in both the Congressional appropriations bill as a requested topic (See: H. Rept. 114-699) and as a part of an RFA (See: https://goo.gl/zQhmN6, accessed July 12, 2017). But notably, the RFA in which a vaccine for Zika is requested is actually a broader request for "Countermeasures Against Select Pathogens," to include a large number of antimicrobial-resistant bacteria or emerging viral pathogens.

On the other hand, the programmatic reasons cited by NIH staff revolve largely around targeting unobservables that are, if anything, likely negatively correlated with scientists' preferences. The staff repeatedly referenced how RFAs are developed to fill "portfolio gaps," or in other words, areas of science where the NIH did not have active grants. These intentions are mirrored in this remark from Thomas Insel, the director of the National Institute of Mental Health (NIMH), who describes the purpose of RFAs as follows[4]:

> "The NIMH uses RFAs to [1] focus on innovation and high-risk science that may suffer in peer review of unsolicited applications ... [2] open up fields that have been relatively neglected ... [3] develop specific, integrated programs that may not be created via unsolicited grants"

Each of these goals revolve around identifying underserved aspects of science, therefore, areas of science relatively less preferred by scientists. This would suggest a negative correlation between any underling trends and the use of RFAs. Whether this correlation is meaningful enough to be empirically relevant is explored throughout the paper.

A.3 Other Relevant NIH Policies

One difference of note regarding RFA and open applications is that after the review and funding decision open applications that fail may be revised and resubmitted again as an open application at a future date. Conversely, the first RFA award decisions are final, so applications may be revised and resubmitted but as an open application. However, because applications initially submitted to RFAs and then resubmitted as an open application are not linked in the NIH data it would be difficult to accurately track such applications. But for the purposes of the following analyses, because this option value of applications is equivalent I only examine outcomes for the first (new) application, and note that whether the option is valued by scientists will not introduce any bias.

Another policy of note is that both successful RFA and open grants may re-apply for continued funding after the initial funding duration expires, but I focus only on first-year direct

[4]Excerpted from: https://goo.gl/2zfPru, accessed July 12, 2017.

costs because this feature applies to both mechanisms, and examining the lifetime value of grants potentially introduces selection concerns as more successful projects are more likely to both pursue these continuation grants and receive them.

In addition to RFAs, NIH also releases "Program Announcements" to solicit certain types of science. However, these calls are not accompanied by set-aside funds made specifically available for competition and in practice vary widely in their format. These announcements are used to facilitate efforts beyond traditional research projects such as conferences, training grants, and other integrated efforts.

There is one variant of the program announcements (Program Announcements with Set-aside funds (PASs)) that I include as a part of the RFA set because for all intents and purposes these mechanisms behave exactly the same as RFAs. This fact was confirmed in discussions with NIH staff. I focus my analyses on RFAs (and PASs) because of their well-defined properties both in terms of scientific scope and set-aside funding.

A.4 RFA Award Process

Once applications to an RFA are received, the review and award process are roughly as follows: (1) applications are reviewed by the designated review group convened specifically for the RFA, where each application either (a) receives a final review score based on the average scores of the group, or (b) receives no score because it is deemed too low quality to be discussed by the entire group; (2) applications are sorted based on this review score; (3) awards are given to the highest ranking applications based on a budget constraint, but not necessarily in perfect rank-order.

Two important features of the award process are of note: the size of the budget constraint and the imperfect rank-order funding. First, it is important to clarify how the budget constraint arises, as it is a key determinant of where the funding threshold is located in the distribution of applications (and in the regression discontinuity design it must be the case that this threshold is randomly located). The largest determinant of this constraint is the amount of funds initially set aside for the RFA, the Purse_j variable. The prior section gave evidence that this amount of funds does not appear to be correlated with anything meaningful in terms of scientists' preferences, which is in line with NIH staff anecdotes that these budgets are influenced by broad scale Institute-wide, unpredictable events. For example, each year an Institute outlays a specific amount of funds for the open competitions, depending on the accuracy of these outlays funds may be substituted away from or towards RFAs. And to the point of imperfect rank-order funding, this arises partly because the NIH generally prefers not to award grants to very scientifically similar applications within an RFA.

However, in both of these instances, the NIH staff is able to influence the award process by (1) endogenously placing the threshold at a certain point to include (or exclude) applications, for example, by withholding some portion of the purse size announced, or lobbying administration for a larger purse, or (2) skipping certain high-rank applications before the budget constraint binds. If either of these events do occur, and they do so for reasons related to the publication potential of the projects, it may bias the treatment effect estimated in

the regression discontinuity model. To address these concerns I present a range of estimates that use varying degrees of cutoffs for the bandwidth of applications included (away from the threshold) and excluded (around the threshold).

A.5 Example RFA

The following is excerpted from the RFA at the following link: https://goo.gl/peirW6, accessed July 12, 2017. For the list of currently active NIH RFA's and other funding opportunities, visit https://goo.gl/hks3K4.

Figure A.1: Development of New Technologies Needed for Studying the Human Microbiome

Department of Health and Human Services

Participating Organizations
National Institutes of Health (NIH), (http://www.nih.gov)

Components of Participating Organizations
This Funding Opportunity Announcement (FOA) is developed as an NIH Roadmap initiative (http://nihroadmap.nih.gov) through the Office of Strategic Coordination (http://dpcpsi.nih.gov/osc/). All NIH Institutes and Centers participate in Roadmap initiatives. This FOA will be administered by the National Human Genome Research Institute (http://www.nhgri.nih.gov) on behalf of the NIH.

Title: Development of New Technologies Needed for Studying the Human Microbiome (R01)

Announcement Type
This Funding Opportunity Announcement (FOA) is a reissue of RFA-RM-08-026.

Request for Applications (RFA) Number: RFA-RM-09-008

Key Dates
Release/Posted Date: July 16, 2009
Opening Date: August 14, 2009 (Earliest date an application may be submitted to Grants.gov)
Letters of Intent Receipt Date(s): August 17, 2009
NOTE: On-time submission requires that applications be successfully submitted to Grants.gov no later than 5:00 p.m. local time (of the applicant institution/organization).
Application Due Date(s): September 14, 2009
Peer Review Date(s): February-March 2010
Council Review Date(s): May 2010
Earliest Anticipated Start Date(s): July 2010
Additional Information To Be Available Date (Activation Date): Not Applicable
Expiration Date: September 15, 2009

Executive Summary

- **Purpose.** The purpose of this FOA is to solicit applications to develop new and improved technologies for obtaining samples of individual microbial isolates or strains, from the human microbiota, suitable for complete genomic sequence analysis. The goal is to expand the number of "reference" microbial genome sequences, which in turn will aid in the analysis of the complex microbial populations resident in and on the human body.
- **Mechanism of Support.** This FOA will utilize the NIH Research Project Grant (R01) grant mechanism and runs in parallel with a FOA of identical scientific scope, RFA-RM-09-009 that solicits applications under the R21 mechanism.
- **Funds Available and Anticipated Number of Awards.** $2 million is available in FY10 for this FOA and the parallel R21 FOA in combination. It is anticipated that 2-4 R01 grants (of duration up to 3 years) and 2-6 R21 grants will be awarded. . Awards issued under this FOA are contingent upon the availability of funds and the submission of a sufficient number of meritorious applications.
- **Budget and Project Period.** Because the nature and scope of the proposed research will vary from application to application, it is anticipated that the size and duration of each award will also vary. Applicants for R01 grants may request a project period of up to 3 years.

RESEARCH SCOPE: The interpretation of metagenomic sequence data is greatly aided by comparison to the genomic sequence of isolated species and genetically different strains of the same species. Yet, only a small proportion of the microbial species resident in or on the human body has been isolated and sequenced. The purpose of this FOA is to support the development of technologies that will allow the determination of the complete, individual genome sequences of substantial numbers of previously uncharacterized members of the human microbiota, to aid in the interpretation of metagenomic datasets obtained from sampling the human body. The following list, which is certainly incomplete, presents examples of strategies that would be supportable under this FOA:

- Development of methods to isolate single microbial cells. These methods would enable the identification, analysis and isolation of individual cells in the human microbiota that satisfy a specified set of criteria.
- New approaches to obtain pure cultures or simple mixed cultures of small numbers of previously uncultivated species would advance the objective of genomic analysis of the human microbiota. Proposed methods that can be applied to a large number of species rather than to any one particular species will take high priority.
- Development, optimization and validation of methods to isolate, amplify, or clone unamplified or amplified DNA of whole genomes from individual cells at high fidelity (e.g., complete coverage, low bias, low chimerism).
- Development of methods to "normalize" the complexity of the population, at either the cellular or DNA level. Such methods would facilitate either the ability to isolate single cells that are rare within a population, or to perform bioinformatics analysis on metagenomic sequences (e.g., by improving the representation of "rare" members).
- Development of methods to enrich the cells of a given species to essential purity. This is the inverse of reducing redundancy, and might be most effective for species whose abundance is already high. Such methods might substitute, at least for DNA sequencing studies, for the ability to establish pure cultures.
- Development of methods that (as a prelude to isolating single microbial cells, or conducting enrichment or normalization) disaggregate cells from the complex mixtures of microbial cells, human cells, and extracellular materials (e.g., biofilms) that comprise human microbial samples. Methods for cell disaggregation should be developed in conjunction with associated methods such as those described above.

Appendix B Medical Subject Heading Indexing & Data Construction

Medical Subject Heading (MeSH) terms comprise the National Library of Medicine's (NLM) hierarchical dictionary. They provide a useful way to classify "types" of science and generate units of observation fit for econometric analyses (cf. Azoulay et al. 2010). Details on the construction and maintenance of MeSH is available at https://www.nlm.nih.gov/mesh/.

The process by which the NLM assigns MeSH terms to documents includes both machine and human review. The algorithm underling the machine assignment step is publicly available (https://ii.nlm.nih.gov/MTI/), with an interactive version of the software available as well (https://meshb.nlm.nih.gov/MeSHonDemand). This indexer can extract the MeSH terms relevant to a body of biomedical text, effectively classifying each abstract into a set of discrete types of science. Using the RFA announcement from the previous section as example, the MTI identified the following MeSH terms as relevant: "Computational Biology", "Human Body", "Chimerism", "Industrial Development", "Biofilms', "Goals', "Genomics', "Metagenomics', "Sequence Analysis, DNA', "Genome', "DNA', "Cell Separation', "Microbiota', "Bias', and "Complex Mixtures". These terms very intuitively capture the goals of this particular RFA.

The MTI faired well with the abstract of this particular paper. It identifies the following terms as relevant: "National Institutes of Health (U.S.)", "Biomedical Research", "Financial Support", "Financial Management" and "Elasticity". Although the MeSH term for elasticity refers not to the economic concept but the mechanical process of resistance and recovery, illustrating the limitations of generalizing this tool.

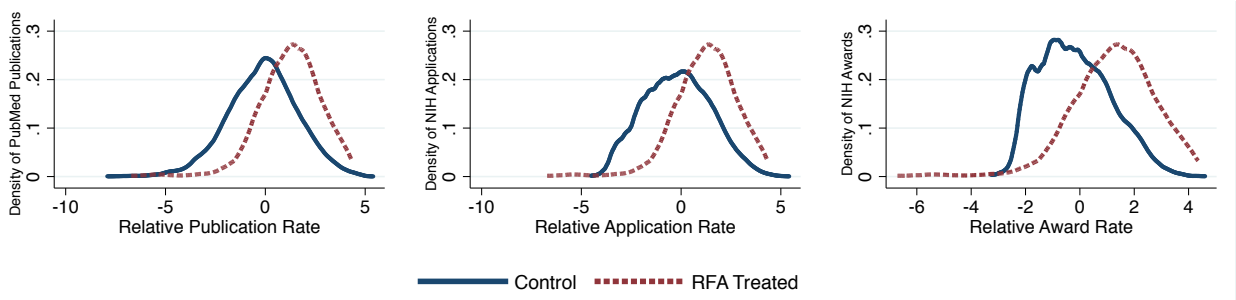
In the analyses using MeSH terms, I control for each MeSH term's position within the MeSH hierarchy using a set of dummy variables that describe each terms' distance (in terms of number of nodes) from its respective top node. Furthermore, I interact this metric with an indicator for each of the seven top nodes to allow this effect to vary within each major set. This eliminates variation that arises simply because certain terms are broader than others (i.e., "Neoplasms," which is 1 nodes from the top node, versus "Large Granular Lymphocytic Leukemia," which is 6 nodes from the top).

For the analyses of Section 2.4, I restrict the sample to include only MeSH terms from the seven major categories that cover purely "scientific" topics: Anatomy; Organisms; Diseases; Chemicals and Drugs; Analytical, Diagnostic and Therapeutic Techniques, and Equipment; Psychiatry and Psychology; Phenomena and Processes.[5] I then include MeSH terms that (1) appear at least once in the NIH application data, (2) occur no more than once in an RFA. Criterion (1) ensures that I examine only MeSH terms at real risk of being pursued by NIH applicants and criterion (2) eliminate any variation in the data that may arise from repeated treatments over the time period I examine.

[5]The other major categories are: Disciplines and Occupations; Anthropology, Education, Sociology, and Social Phenomena; Technology, Industry, and Agriculture; Humanities; Information Science; Named Groups; Health Care; Publication Characteristics; Geographicals.

The cross-sectional distributions of treated and control terms is presented in Figure B.2.

Figure B.2: Distribution of Scientific Subjects, Treated & Control



(a) PubMed Publications **(b)** NIH Applications **(c)** NIH Awards

Plots the average log number of abstracts related to a MeSH term, after controlling for the term's position within the MeSH hierarchy. Abstracts are from the PubMed database of biomedical journal articles (Panel a), all NIH grant applications (Panel b) or successful NIH applications (Panel c).

Appendix C Direction, Similarity & the *pmra* Algorithm

C.1 Choice of Direction & the Role of Similarity

Broadly speaking, there are two main reasons why redirections would be difficult in this setting: they require both tangible (i.e., equipment) and intangible costs (i.e., preferences, cognition). With this in mind, the elasticity I estimate is a *behavioral parameter* in that its magnitude is driven by these tangible and intangible aspects.

In the biomedical sciences, the pecuniary costs of adjustment are substantial; individual pieces of lab equipment routinely cost in excess $100,000. And human capital theory has long appreciated the limitations of specialized knowledge (e.g., Becker 1962), with much emphasis placed on the potential for one's prior endeavors to shape and constrain their search and evaluation of new ideas (Nelson and Winter 1982; Gavetti and Levinthal 2000; Boudreau et al. 2016). Beyond any potential cognitive constraints, scientists have been seen to exhibit preferences over the nature of their work (Stern 2004), and be influenced by social forces (Stuart and Ding 2006; Ding et al. 2006). But notably, many of these studies focus on discrete changes in direction often primarily related to commercialization activities, and not directional adjustments in general. Certainly the commercial transition is one of obvious economic impact, but results on the decisions of direction *before* commercialization or with regards to the *type of science* have been very limited to date. Bhattacharya and Packalen (2011) examine the direction of basic science more broadly using the occurrence of biomedical terms in publications to classify the direction of science as a whole, and find that in the aggregate biomedical scientists do appear to pursue fields related to diseases with higher prevalence as well as those with an increasing underlying fertility.

Two studies that examine the movement of scientists across fields use journal article retractions (Azoulay et al. 2015a) and untimely deaths (Azoulay et al. 2016) as shocks to individuals and fields, respectively. Both papers use the algorithm described below to estimate the degree of scientific redirection, although their implementation relies on publications which means that only successful (per publications) redirection is observed (and the latter is also true of Bhattacharya and Packalen's analyses (2011)). The authors find that following these events, which essentially remove barriers to operating in a particular type of science (e.g., lower competition or fewer "gate-keepers"), scientists from neighboring fields enter. My paper builds on this literature by estimating the costs of scientific redirection in general and before outcomes are realized.

C.2 *pmra* Algorithm Details

Lin and Wilbur (2007) develop a topic-based similarity model based on Bayes' Theorem that estimates the probability that an individual is interested in document a given expressed interest in document b, or in other words, what is the likelihood that a and b are scientifically similar. They focus on the following relationship:

$$\Pr(a|b) \propto \sum_{j=1}^{N} \Pr(a|s_j) \Pr(b|s_j) \Pr(s_j),$$

where $\{s_1, ..., s_N\}$ denotes the entire set of mutually exclusive topics that could possibly be contained within a, b, or any other document of interest. Lin and Wilbur (2007) then make assumptions about the underlying arrival rates of terms within documents (Poisson) and how likely the occurrence of a term within a document actually reflects the true nature of that document. From these assumptions, the authors arrive at a topic weighting function, $w_{j,x}$, that describes that how important a topic s_j is to any document x, and a document scoring function, $Sim(a, b)$, that quantifies the similarity between a and b, given by:

$$w_{j,x} = \lambda_{j,x} \times \sqrt{\frac{1}{f_j}}$$

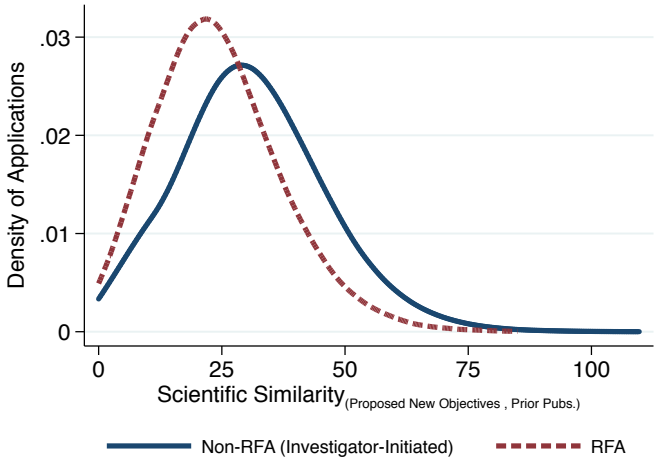
$$Sim(a, b) = \sum_{j=1}^{N} w_{j,a} \times w_{j,b},$$

where f_j is the frequency that topic s_j occurs in the universe and $\lambda_{j,x}$ is based on a series of Poisson arrival rate parameters and the number of times that topic s_j occurs in document x. Intuitively, two documents are more likely to be similar when they both use topics that are rare ($1/f_j \uparrow$) many times ($\lambda_{j,x} \uparrow$). The authors estimate, optimize and experimentally confirm parameters within $\lambda_{j,x}$ to align with human assessments. Loosely speaking, this approach is analogous to the cosine similarity approach previously used to estimate scientific similarity (i.e., Boudreau et al. (2016)), here, weighted by the rarity of intersecting topics.

For specific details on the algorithm and how topics are defined, see Lin and Wilbur (2007), and for a broader overview of how this algorithm is implemented at the National Library of Medicine, see https://goo.gl/PbvpvW, accessed July 12, 2017.

C.3 Full Application Similarity Distribution, RFA v. Open

Figure C.3: Distribution of Application-Scientist Similarity Scores



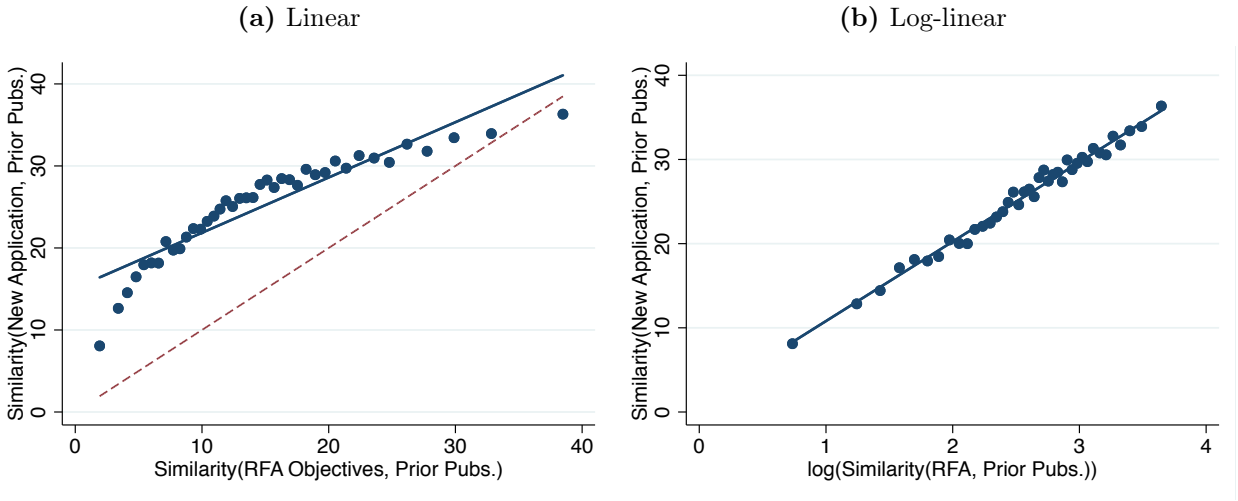
Includes all new, R01 applications.

C.4 Entry Model Similarity Adjustment

Figure C.4 plots the scientist-application similarity as a function of the scientist-RFA similarity. If the first metric (observable for entrants only) is equivalent to the second (observable for all), then there should be a 1:1 relationship between these variables and each data point would lie on the 45 degree line. However, it is clear in Figure C.4 Panel (a) that this is not the case. It appears that relative (percentage) increases in the scientist-RFA similarity implies level (absolute) increases in the scientist-application similarity.

Figure C.4 Panel (b) uses a log-transformation of scientist-RFA metric to explore this log-linear relationship, which fits the data very well. Given the good fit, I predict scientist-application scores using this log-linear model. This predicted value represents, on average, how large of a redirection each scientist would require to enter an RFA given their observed scientist-RFA similarity score. This predicted value represents, on average, how large of a redirection each scientist would require to enter an RFA given their observed scientist-RFA similarity score.

Figure C.4: Relationship between Similarity Scores for Entrants

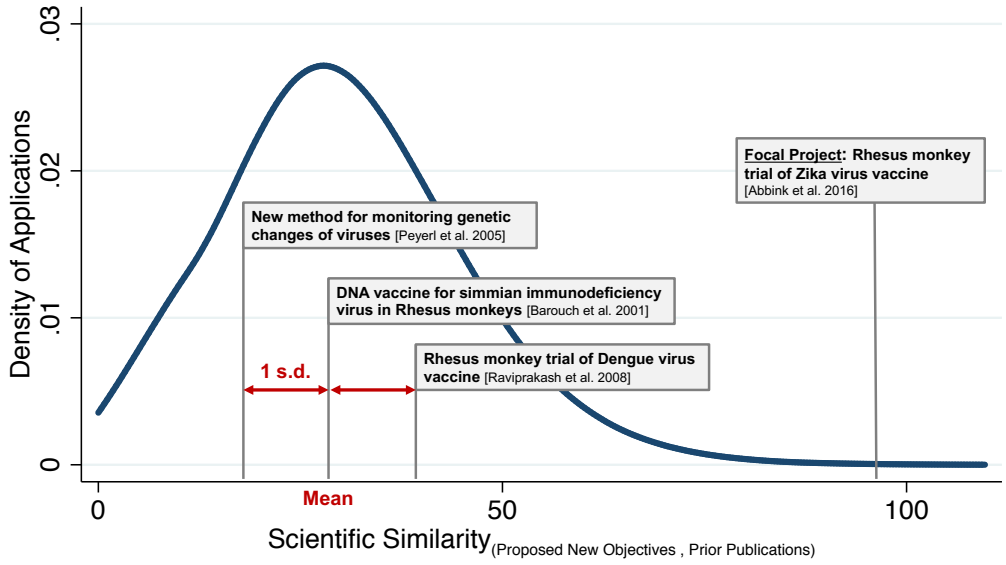


Note: Binned scatterplots ($N_{bins} = 40$) of *pmra*-generated similarity scores for scientists that enter an RFA with fitted lines in solid, and dashed 45°. Plots the "distance-needed" (x-axis) and the "distance-traveled" (y-axis).

C.5 Qualitative Interpretation

The following two figures (C.5 and C.6) plot the empirical distribution and statistics of *pmra* scores from NIH applications, based on scientists prior publications and their new application. The example figures are generated to consider a publication from the biomedical (C.5) or economics literatures (C.6) as a focal project, e.g., a scientist's most recent work, and ask what would that scientist's next project look like given the NIH sample moments. The focal projects are plotted at the average *pmra* score for documents scored against itself, with three expected "next projects" plotted at the NIH sample mean and a +/− one standard deviation increase (more similar) and decrease (less similar) using those publications approximate *pmra* scores relative to the focal publication.

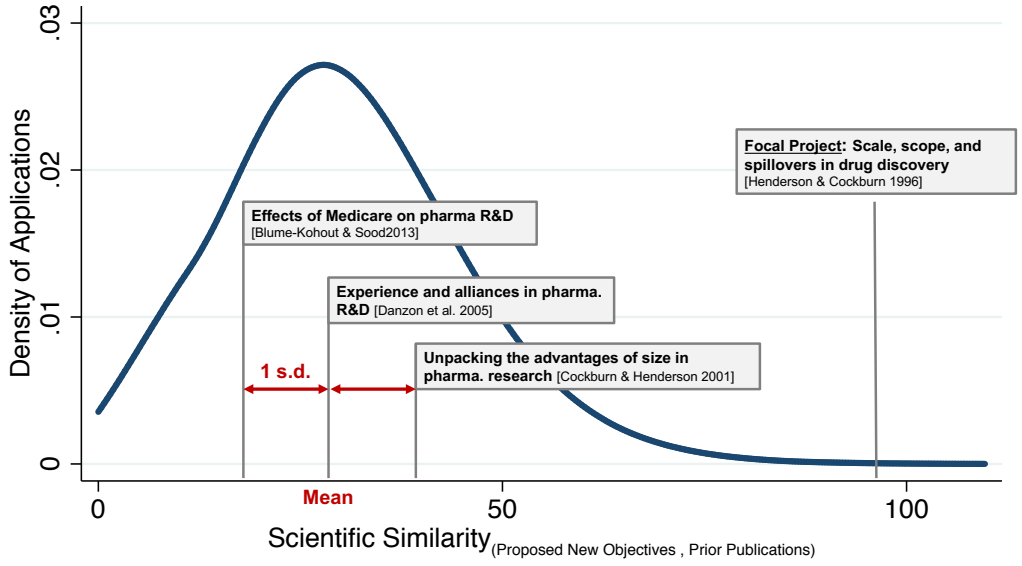
Figure C.5: *pmra* Distribution: Biomedical Examples



Citations:

1. Abbink et al. (2016). Protective Efficacy of Multiple Vaccine Platforms Against Zika Virus Challenge in Rhesus Monkeys. *Science*, 353(6304): 1129?1132.
2. Raviprakash et al. (2008). A tetravalent dengue vaccine based on a complex adenovirus vector provides significant protection in rhesus monkeys against all four serotypes of dengue virus. *Journal of Virology*, 82(14):6927-6934.
3. Barouch et al. (2001). Elicitation of high-frequency cytotoxic T-lymphocyte responses against both dominant and subdominant simian-human immunodeficiency virus epitopes by DNA vaccination of rhesus monkeys. *Journal of Virology*, 75(5):2462-2467.
4. Peyerl et al. (2005). Use of molecular beacons for rapid, real-time, quantitative monitoring of cytotoxic T-lymphocyte epitope mutations in simian immunodeficiency virus. *Journal of Clinical Microbiology*, 43(9):4773-4779.

Figure C.6: *pmra* Distribution: Economics Examples



Citations:

1. Henderson & Cockburn (1996). Scale, scope, and spillovers: the determinants of research productivity in drug discovery. *The RAND Journal of Economics*, 27(1): 32-59.
2. Cockburn & Henderson (2001). Scale and scope in drug development: unpacking the advantages of size in pharmaceutical research. *Journal of Health Economics*, 20(6): 1033-1057.
3. Danzon et al. (2005) Productivity in pharmaceutical-biotechnology R&D: the role of experience and alliances. *Journal of Health Economics*, 24(2): 317-339.
4. Blume-Kohout & Sood (2013). Market size and innovation: Effects of Medicare Part D on pharmaceutical research and development. *Journal of Public Economics*, 97: 327-336.

Appendix D Motivating Theory

To motivate the analyses, consider a world where each scientist $i = \{1, 2, ..., N\}$ can costlessly adjust the trajectories of their research. In this world, they receive the outside option of 0, or can compete for research funding by submitting proposals to one of two competitions $j = \{1, 2\}$, where each a submission's quality is based on scientists' random draws from i.i.d. quality distributions.

In the competitions, funds are allocated based on an award function W, which is well-behaved and maps three variables into awards: (1) each submission's quality q_{ij}; (2) the vector of other submissions $\mathbf{q_j}$, where $n_j \leq (N-1)$ is the number of each submission's competition; (3) the total funds available P_j. Realized awards are then given by $w_{ij} = W(q_{ij}, \mathbf{q_j}, P_j)$. It is assumed that $\frac{\partial W}{\partial q_{ij}} > 0$ and $\frac{\partial W}{\partial N_j} < 0$; higher quality applications in less contested competitions perform better. Furthermore, assume that each individual is extremely small relative to the full set of potential entrants, so although individuals may form strategic expectations about n_j, the general equilibrium effects are negligible; scientists are atomistic.[6]

Let the expected payoff from entry into j simply be $V_{ij} = \mathbb{E}[w_{ij}] - c$, where entry costs c are constant across scientists and competitions. In this case, scientists enter $j = 1$ if three conditions hold: (1) $V_{i1} > 0$, (2) $\mathbb{E}[w_{i1}] > c$, and (3) $\mathbb{E}[w_{i1}] > \mathbb{E}[w_{i2}]$. Clearly, in equilibrium both options, regardless of the amount of funds are made available, should see competition to the point that their expected values are equivalent. If the payoff of entering one of the competitions is larger, then the "free-range" scientists will simply enter and compete down the expected value until it equates with the alternative. This world embodies the zero-profit nature of perfectly competitive markets with free entry.

If the entry costs to one of the competitions is increased, say $c_1 > c_2$, then in equilibrium $\mathbb{E}[w_{i1}] > \mathbb{E}[w_{i2}]$ and most relevant for the analysis, $c_1 - c_2 = \mathbb{E}[w_{i1}] - \mathbb{E}[w_{i2}]$. This implication is a staple of traditional industrial organization models where markets with higher fixed costs also have larger profit margins. And thus, with unbiased estimates of expected payoffs in hand, one can compute the difference in entry costs.

More specifically though, in this setting I am interested in estimating how a certain feature of these competitions - the degree to which they require scientists to adjust their work - might influence entry costs, and therefore, create a wedge in expected payoffs. Rewrite V_{ij} to now be $\mathbb{E}[w_{ij}] - c(s_{ij}, \xi_j)$ with s_{ij} describing the similarity between individual i's scientific expertise and the type of science required for entry into j, and ξ_j capturing the (potentially zero) fixed value all scientists place on j. First, assume that only this similarity factor influence costs and thus, $\xi_j = 0$. Then it follows that on average $c(\widehat{s_{i1}}) - c(\widehat{s_{i2}}) = \widehat{\mathbb{E}[w_{i1}]} - \widehat{\mathbb{E}[w_{i2}]}$, where $\widehat{x}$ denotes the average of variable x across all i. This relationship implies that $\frac{\partial V_j}{\partial \widehat{s_{ij}}} = \frac{\partial V_j}{\partial \widehat{\mathbb{E}[w_{ij}]}}$, providing a way to relate the average marginal gains in expected awards to average marginal differences in similarity - precisely the elasticity of interest: $\frac{\partial \widehat{s_{ij}}}{\partial \widehat{\mathbb{E}[w_{ij}]}}$, or how large of a change in science can be induced by a given change in funds?

[6] In the sample used for this analysis, there are roughly 16,000 unique applicants to the NIH per year.

Taken together, if $\xi_j = 0$, and I can empirically estimate (1) the difference in expected awards between the RFA and open mechanisms that arises from the NIH's exogenous allocation decisions and scientists' endogenous responses ($\widehat{\partial \mathbb{E}[w_{ij}]}$) - the "RFA premium" - as well as (2) the average level of redirection that RFAs induce beyond what is observed in the open applications ($\widehat{\partial s_{ij}}$), then I can identify the elasticity.

However, if $\xi_j \neq 0$, then I will instead be estimating $c(\widehat{s_{i1}}, \xi_1) - c(\widehat{s_{i2}}, \xi_2)$ and conflate the costs of changes in s with some fixed costs (or benefits) captured by ξ_j. In this case I can only clearly estimate the RFA premium ($\widehat{\mathbb{E}[w_{i1}]} - \widehat{\mathbb{E}[w_{i2}]}$), and will overestimate (underestimate) the elasticity if $\xi_j > 0$ ($\xi_j < 0$).

Appendix E Robustness Tests & Extensions

E.1 Entry Model Heterogeneity Analyses, Full Tables

Table E.1: Non-linear Elasticity across the Similarity Distribution

	(1)	(2)	(3)	(4)	(5)
Scientific Similarity	0.125*	0.304**	1.19***	3.37***	54.9***
	(0.070)	(0.139)	(0.214)	(0.373)	(2.230)
Purse	0.736***	1.32***	1.61***	2.49***	10.5***
	(0.206)	(0.472)	(0.456)	(0.455)	(0.766)
Instrumented Competition	-0.262**	0.0217	-0.766***	-1.69***	-12.5***
	(0.122)	(0.267)	(0.235)	(0.290)	(0.875)
$N_{scientist-RFA}$	14,235,537	7,760,802	11,015,939	11,033,012	11,033,461
RFA controls	Y	Y	Y	Y	Y
Scientist F.E.	Y	Y	Y	Y	Y
Similarity Quintile	1 (least)	2	3	4	5 (most)
ε	2.254	0.641	0.085	0.035	0.020

Note: Independent variable is a dummy indicating entry. All dependent variables are standardized, and all coefficients and standard errors are scaled by 10^{-4}. mean(Entry$_{ij}$) = 2.68×10^{-4}. N_{RFA} 394. RFA controls are a vector of characteristics including dummies for the fiscal year, the NIH Institute administering the RFA. *p<0.1 **p<0.05 ***p<0.01; standard errors clustered within RFAs. ε is the elasticity of science at the sample means/s.d.s per the coefficients on Scientific Similarity and Purse.

Table E.2: Elasticity and Experience, by Quintile of Year of First Publication

	(1)	(2)	(3)	(4)	(5)
Scientific Similarity	13.5***	12.2***	7.93***	3.65***	1.63***
	(0.635)	(0.516)	(0.358)	(0.225)	(0.161)
Purse	5.64***	5.36***	4.03***	2.31***	0.929***
	(0.520)	(0.528)	(0.454)	(0.411)	(0.192)
Instrumented Competition	-5.71***	-5.32***	-3.98***	-2.00***	-0.792***
	(0.409)	(0.359)	(0.313)	(0.236)	(0.016)
$N_{scientist-RFA}$	9,572,028	9,843,386	10,550,538	9,333,094	7,821,904
RFA controls	Y	Y	Y	Y	Y
Scientist F.E.	Y	Y	Y	Y	Y
Avg. Year of 1^{st} Pub.	1974	1988	1995	2000	2004
ε	0.110	0.123	0.162	0.260	0.295

Note: Independent variable is a dummy indicating entry. All dependent variables are standardized, and all coefficients and standard errors are scaled by 10^{-4}. mean(Entry$_{ij}$) = 2.68×10^{-4}. N_{RFA} 394. RFA controls are a vector of characteristics including dummies for the fiscal year, the NIH Institute administering the RFA. *p<0.1 **p<0.05 ***p<0.01; standard errors clustered within RFAs. ε is the elasticity of science at the sample means/s.d.s per the coefficients on Scientific Similarity and Purse.

E.2 Fuzzy RD Sample Summary Statistics

Table E.3: Regression Discontinuity Sample Summary Statistics

	Type	Mean	s.d.	Median	N
Panel A: Application-level					
Win	{0,1}	0.379	0.485	0	2672
Review Score	Cont.	230.7	62.95	223	2672
Applicant has M.D.	{0,1}	0.398	0.490	0	2672
Applicant has Ph.D.	{0,1}	0.723	0.448	1	2672
Multiple P.I.s	(0,1)	1.000	0.019	1	2672
Study involves animals	{0,1}	0.456	0.498	0	2672
Study involves humans	{0,1}	0.609	0.488	1	2672
All-year Total Funds Requested ($M)	Cont.	1.733	1.095	1.564	2672
Pre, Pubs.	Count	24.21	22.79	20	2672
Pre, Pubs., F/L Author	Count	15.21	15.72	11	2672
Post, Pubs.	Count	12.35	13.78	8	2672
Post, Pubs., F/L Author	Count	6.875	8.209	4	2672
Pre, Avg. Similarity	Cont.	6.783	4.504	6.148	2366
Post, Avg. Similarity	Cont.	8.665	5.451	8.135	2299
Diff., Avg. Similarity (%)	Cont.	0.236	0.429	0.211	2037
Panel B: RFA-level					
Scored Applications	Count	19.04	12.63	17	141
Realized Purse (1^{st}-year Direct Costs, $M)	Cont.	1.946	1.384	1.533	141
Realized Purse (1^{st}-year Total Costs, $M)	Cont.	2.549	2.013	2.081	141
Pr(Win)	[0,1]	0.420	0.169	0.400	141
Avg. Award Size (All-year Total Costs, $M)	Cont.	1.928	1.066	1.757	141
Avg. Award Length (years)	Count	4.678	1.269	4.667	141
Year of RFA	Count	2003.8	1.957	2004	141

Note: Sample includes applications to RFAs held from 2001 to 2007, conditional on receiving a review score. Multiple P.I.s refers to applications with more than 1 principle investigator listed. For these applications, outcomes include publications by all authors involved. F/L indicates that the focal applicant was listed as the first or last author on the publication. Similarity scores are measures by the *pmra* algorithm as the similarity between each publication and the research objectives section of the RFA applied to. Differences reported are post outcomes minus pre. Some observations are dropped for particular similarity outcomes because these measures require at least one publication to occur.

E.3 Main Results Robustness

Table E.4: Robustness Tests: Entry Model

	With Pub. Count Control		log, max($pmra$)		Unadjusted max($pmra$)		Unadjusted max($pmra$, ≤5 yrs.)		mean($pmra$)		median($pmra$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Scientific Similarity	5.63***	6.25***	7.33***	7.52***	9.81***	10.8***	9.06***	9.77***	4.80***	5.36***	4.09***	4.52***
	(0.0383)	(0.246)	(0.351)	(0.350)	(0.389)	(0.412)	(0.358)	(0.375)	(0.223)	(0.234)	(0.205)	(0.210)
Purse	1.40***	2.84***	2.13***	3.00***	1.47***	2.89***	1.44***	2.83***	1.40***	2.97***	1.38***	3.03***
	(0.0249)	(0.387)	(0.458)	(0.634)	(0.425)	(0.385)	(0.424)	(0.385)	(0.423)	(0.392)	(0.418)	(0.409)
Instrumented Competition		-2.38***		-1.69**		-2.84***		-2.61***		-2.50***		-2.53***
		(0.216)		(0.788)		(0.161)		(0.163)		(0.260)		(0.331)
ε	0.107	0.197	0.290	0.399	0.131	0.234	0.151	0.276	0.113	0.215	0.138	0.275

Note: *p<0.1 **p<0.05 ***p<0.01; standard errors clustered within RFAs. All coefficients and standard errors are scaled by 10^{-4}. mean(Entry) = 2.7×10^{-4}. $N_{scientist-RFA} = 55,099,570$. N_{RFA} 394. Results from OLS estimation of Eq. **??**, potentially including instrumented competition as outlined in Section 3.2. All models include scientist-fixed effects and a vector of RFA-specific characteristics including dummies for the fiscal year, the NIH Institute administering the RFA. ε is the elasticity of science at the sample means per the coefficients on Scientific Similarity and Purse. Cols. 1-2 recreate Table 3 Col. 7, including the scientist's number of publications as a control (with and without competition control). Cols. 3-4 use a log transformation of the main variables (which are standardized in all other columns). Cols. 5-8 use the raw $pmra$ scores, with Cols. 7-8 only including publications that occur within 5 years of the focal RFA. Cols. 9-10 use the average $pmra$ scores for each scientist's set of publications scored relative to each RFA, and likewise Cols. 11-12 use the median of these scores.

Table E.5: RFA versus Open Application Robustness Tests: Similarity & Award Size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	y : Scientific Similarity				y : Award Size			
1{RFA}	-0.123***	-0.125***	-0.0764***	-0.0604***	0.0824***	0.0798***	0.199***	0.114**
	(0.0147)	(0.0163)	(0.0147)	(0.0160)	(0.0162)	(0.0161)	(0.0401)	(0.0460)
$N_{applications}$	36237	36170	18582	18562	8242	8691	1123	1533
$N_{scientists}$	24864	24809	7408	7400	7445	7810	529	717
mean(y)	12.2	11.1	12.1	10.9	399237	1911925	416968	2010625
Version	avg.	med.	avg.	med.	1^{st} Year Tot. Costs	All Year Tot. Costs	1^{st} Year Tot. Costs	All Year Tot. Costs
Scientist F.E.			Y	Y			Y	Y
Area-Time F.E.	Y	Y	Y	Y	Y	Y	Y	Y

Note: *p<0.1 **p<0.05 ***p<0.01; standard errors clustered within scientists. Estimates of Equation **??** using alternative specifications and transformations. Avg. and Med. scientific similarity refers to taking the average and median *pmra* scores for each scientist's set of prior publications. Total Costs refer to the sum of both Direct and Indirect costs awarded, with All Year costs referring to the total amount of funds awarded over the lifespan of the focal grant award.

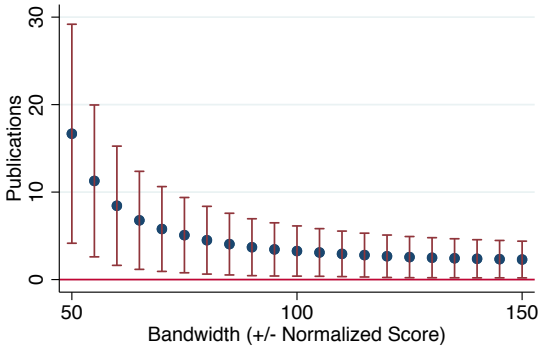
Table E.6: Fuzzy RD Results incl. Pre-Period Control

	(1)	(2)	(3)	(4)	(5)	(6)
	F/L Author Pub Count			Relative Similarity Diff.		
Win	1.167**	3.914**	3.681***	-0.00369	0.381***	0.357***
	(0.483)	(1.768)	(1.143)	(0.0252)	(0.0999)	(0.0746)
N	2408	2408	2408	1822	1822	1822
mean Dep. Var.	7.058	7.058	7.058	0.219	0.219	0.219
RFA F.E.	Y		Y	Y		Y
Spec.	OLS	RD	RD	OLS	RD	RD

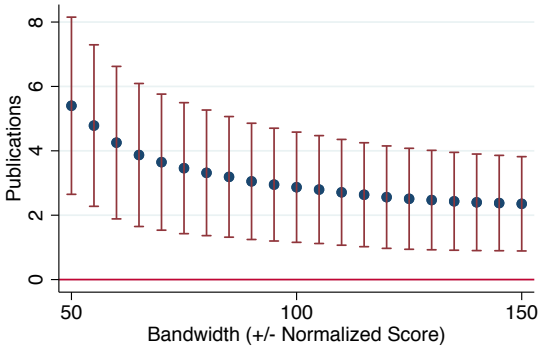
Dependent variables and specifications described in text. Includes a single control variable for the count of publications in the pre period (Cols. 1-3), or the average similarity scores for publications in the pre period (Cols. 4-6). *p<0.1 **p<0.05 ***p<0.01; standard errors clustered within RFAs.

Figure E.7: Fuzzy Regression Discontinuity Results, Bandwidth Varying

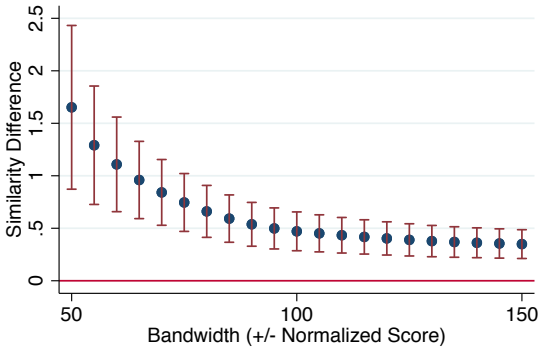
(a) F/L Pubs.: Full Sample



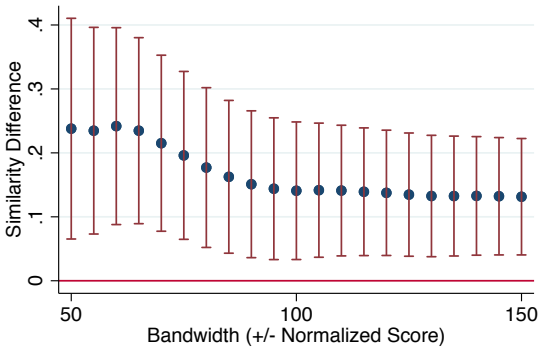
(b) F/L Pubs.: Exclude Border Apps.



(c) Similarity Change: Full Sample



(d) Similarity Change: Exclude Border Apps.



Note: Plots the LATE estimates (1 model per point, with 95% C.I.) from Eqs. **??** using the fuzzy regression design, changing the bandwidth score used to include applications on either side of the threshold. Panels (a) and (c) include all applications; Panel (b) and (d) drop the two applications from each RFA that border the threshold.

Table E.7: Fuzzy RD LATE Heterogeneity

	Publication Count					Similarity Difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	log(cites)	log($\frac{\text{cites}}{\text{pub}}$)	Any > 50%ile	Any > 75%ile	Any > 90%ile	> 50%ile	> 75%ile	> 90%ile
Win	0.0968	-0.107	0.172**	0.236***	0.0821	0.503***	0.349***	0.300***
	(0.288)	(0.202)	(0.0825)	(0.0873)	(0.0862)	(0.119)	(0.0912)	(0.0642)
$N_{\text{applications}}$	1888	1888	2684	2684	2684	1999	1999	1999
mean Dep. Var.	4.739	2.887	0.669	0.569	0.402	0.500	0.251	0.101
Dep. Var. Type	log(Count)	log(Cont.)	$\{0,1\}$	$\{0,1\}$	$\{0,1\}$	$\{0,1\}$	$\{0,1\}$	$\{0,1\}$
%ile Sample	n/a	n/a	PubMed (Citations)			Estimation Sample		
RFA F.E.	Y	Y	Y	Y	Y	Y	Y	Y

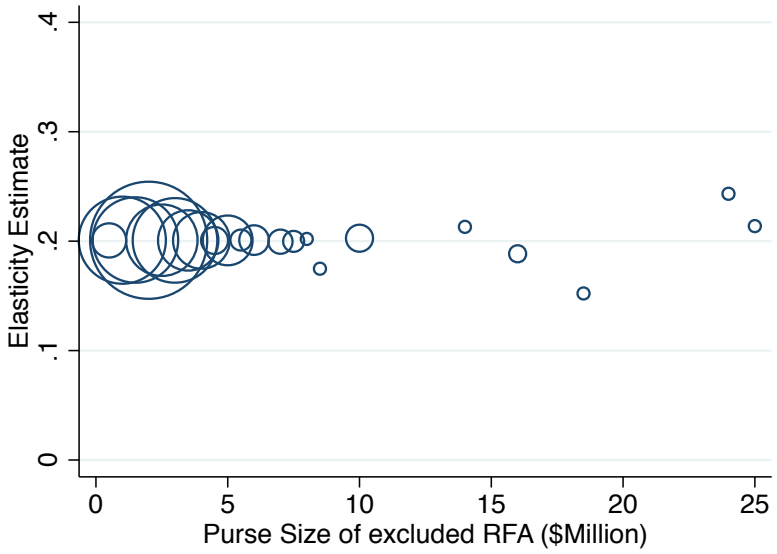
Note: *p<0.1 **p<0.05 ***p<0.01; standard errors clustered within RFAs. For Cols. (1-2), the dependent variable is the log transformed total number of citations (1) or citations per publication (2) post award decision. The dependent variable in Cols. (3-5) is a dummy that equals one when if any of the post award decision applications are above the X^{th} percentile of the distribution of non-self citation counts for all articles in PubMed as of 2015. The dependent variable in Cols. (6-8) is a dummy that equals one when if the relative difference in average similarity scores pre-post is above the X^{th} percentile of the sample distribution.

E.4 Entry Model Identification Tests

RFA-level Jackknife: In order to examine how much the main estimates might be driven by particular RFAs, Figure E.8 plots the range of elasticity estimates from separate regressions where RFAs are dropped one at a time. The range of these estimates is very narrow (mean =0.201, s.d. = 0.004, min = 0.152, max = 0.243). And there is no significant correlation between the elasticity estimate and the purse size of the excluded RFA (corr = 0.024, p-value = 0.6214).

Since Figure E.8 suggests that the largest variability arises when dropping RFAs from the upper end of the purse distribution, I explore how much these particular RFAs are driving the main results. Table E.8 reports the preferred specification of the entry model, estimated without RFAs from the the top 5% of the purse distribution (≈$8M). Dropping these large RFAs does not result in any meaningful change to the main estimate of 0.2. Since purse variation amongst these normal-sized RFAs is much more likely to be driven by exogenous forces such as budget shocks, this consistency is reassuring that the key identification assumptions of this paper are valid.

Figure E.8: Jackknife Elasticity Estimates



Note: Binned scatterplots of elasticity estimates from separate regressions each dropping a single RFA, plotted as a function of the purse size of the excluded RFA. Bins are weighted by the number of RFAs of each purse size, rounded to $500,000.

Implied Value of Competition: Recall, the estimated coefficient on $\widetilde{n_{ik}}$ indicates how entry probabilities change as the number of expected competitors increases. Just as in the case of similarity, I can relate this magnitude to the coefficient on Purse_k to identify the implied dollar value that scientists place on competition ($\frac{F'(\text{Purse}_k)}{D'(\widetilde{n_{ik}})}$); how much larger must a purse size be to make scientists indifferent to the addition of one more expected competitor. But unlike the case of similarity, the magnitude of this particular parameter can be directly

Table E.8: Entry Model & Implied Elasticities, Dropping Largest RFAs

	(1)	(2)
Scientific Similarity	6.296***	7.644***
	(0.243)	(0.354)
Purse	1.572***	1.574***
	(0.122)	(0.253)
Instrumented Competition	-1.747***	-0.988**
	(0.127)	(0.415)
Functional Form	Standardized	log
RFA controls	Y	Y
Scientist F.E.	Y	Y
$\boldsymbol{\varepsilon}$	0.180	0.206

Note: *p<0.1 **p<0.05 ***p<0.01; standard errors clustered within RFAs. Independent variable is the probability of entry. All dependent variables are standardized in parentheses clustered at the RFA level, are scaled by 10^{-4}. mean(Entry) = 2.45×10^{-4}. $N_{scientist-RFA} = 56,241,530$. N_{RFA}=373. RFA controls are a vector of characteristics including dummies for the fiscal year, the NIH Institute administering the RFA. Excludes RFAs in the top 5% of the Purse distribution (21 of possible 394). Functional form indicates whether all independent variables are either standardized or log-transformed. $\boldsymbol{\varepsilon}$ is the elasticity of science at the sample means per the coefficients on Scientific Similarity and Purse

estimated from the data under some reasonable assumptions. If RFA purse sizes are truly exogenous (i.e., $\mathbb{E}[\xi_{ik}|\text{Purse}_k = 0)$, then the dollar value of competition implied by the model should be equal to the estimate directly taken from the data.

How can I directly estimate this parameter? Because the purse is held fixed, the addition of one more competitor ($\partial\widetilde{n_{ik}} = 1$) inherently decreases the expected value of the purse for others by an amount equal to that competitor's expected value, which is simply (1) the average marginal win probability times (2) the average award size. I previously estimated the average marginal win probability as 0.19934 (s.e. = 0.00775) using OLS with RFA-fixed effects. This estimate is virtually identical to the average win probability conditional on entry (0.19932), which suggests that the marginal entrants fair just as well as the average entrant.[7] I cannot directly estimate the average marginal award size in this context (because winning is defined as receiving an award). However, given the previous result, I assume that just as for win probabilities, the average and marginal awards are equivalent in size, the sample average of $407K.

Now, I can directly estimate the decrease in expected payoff give one additional entrant - it

[7]The typical economic intuition is that marginal entrants to any market will be less competitive - new entrants tend to come from lower points on the quality distribution. But recall that here, the number of potential entrants is upwards of 140,000. So even the 25^{th} entrant will likely be coming from very close on the distribution to the first 24.

equals $0.19 \times 407{,}000 = 77{,}330$.[8] The estimates from Table **??** indicate that, at the sample means, scientists are indifferent between competing with 1 additional entrant and about a \$75,000-\$78,000 decrease in purse size (per Columns 5-7). The dollar-value of competition implied by the entry model aligns closely with the outcomes observed. This would not be the case if $\mathbb{E}[\xi_{ik}|\text{Purse}_k] \neq 0$. Again, there is no evidence that the amount of funds made available for any given RFA is meaningfully correlated with any features of the science targeted within.

E.5 Back of the Envelope: Adjustment versus Production

The estimates thus far are helpful in thinking about the relative costs of incentivizing adjustments, but it would be useful to better understand the absolute magnitude of their importance. One way to begin to answer this question is to ask a more specific question: what share of the costs observed in this setting scale with adjustments relative to production?

Simply put, each scientific endeavor has costs associated with both the intended *rate* (i.e., the size) and *direction* of output. Every new experiment costs money, and new experiments in new directions cost more money. Understanding the division of these costs is essential for managers and policymakers who have the capability to selectively allocate funds to these activities.

To get a sense of this division of costs, I use the following thought experiment: for the average scientist who just received an average NIH grant to conduct the average project, how much of those funds would the scientist be willing to forego in order to instead reproduce their prior, most similar project. That is to say, holding fixed the intended rate of production (e.g., maximize publication output per dollar) how much smaller could this psuedo-replication project be and still have the scientist prefer it. Loosely speaking, the goal here is to estimate the compensating differential that arises purely due to the changes in direction that scientists undertake in their new projects.[9]

We now know that given some replication project ($p = 0$) and average new project ($p = 1$), the two are related via the elasticity of science ε, as given by:

$$\varepsilon = 0.2 = \frac{s_0 - s_1}{s_0} \Big/ \frac{a_1 - a_0}{a_0}, \tag{1}$$

where s_p and a_p are the scientific similarity and award sizes, respectively, and the 0.2 is based on the results of the entry model. This is the relative change in direction that can be induced with some relative change in award. The similarity and award sizes for the average project are simply the sample averages. The unknowns are for the "replication" project. The

[8] Because the Purse_j variable is reported in RFAs in terms of first-year total costs, the award size used for these calculations is also first-year total costs.

[9] In reality, scientists very likely value the inputs and outputs of true replication studies very differently. The point of this exercise is simply identify how much of total project costs scale with changes in similarity, not to predict the actual compensating differential. Thus, I assume a fixed value of funds across levels of similarity.

most reasonable way to infer s_1 from my data is to examine the highest level of similarity observed in the data (the smallest redireciton). The 99^{th} percentile of *pmra* score between a scientist's past publication and new application observed in the data is about 100. Thus, I take this value as my best estimate for the similarity of a project with effectively no change in direction.[10]

Setting $s_0 = 100$ and using the sample averages for s_1 (30) and a_1 ($300K) requires that a_0 equals roughly $66,000. That is to say, my estimates imply that the costs associated with the average level of redirection observed in the data total $234,000 (=$300K-66K). This magnitude suggests that roughly 78% of the costs associated with NIH-sponsored research scale with directional adjustments.

[10]This is also the approximate average score the algorithm generates when computing the similarity between two identical publication abstracts. That there is any variation in these self-scores is based on the probabilistic modeling of the algorithm. The value of 100 is not indicative of any percentage as the *pmra* algorithm generates a likelihood.

Appendix F Counterfactual Methodology

This section outlines the approach to estimating the marginal costs of inducing applications presented in Section 3.6. The challenge is that the focal model consists of linear regressions which are additively separable in the two coefficients used to identify the focal elasticity: the coefficients on the similarity measure and the purse size. Thus, a given elasticity estimate could be obtained through many different of adjustments to either, or both of these coefficients (e.g. $0.2 = \frac{1}{5} = \frac{0.004}{0.02} = \frac{20}{100}$,...).

In order to arrive at new coefficients for predict outcomes at a given elasticity, while maintaining internal consistency to the model, I use the following relationship to project the two new coefficients as a function of a certain elasticity:

$$\varepsilon' = \frac{\partial_{entry}\text{Purse} \times (1+\delta)}{\partial_{entry}\text{Similarity} \times (1-\delta)},$$

where ∂_{entry}Purse and ∂_{entry}Similarity are the changes in entry probability given relative (%) changes in the corresponding variables based on the regression coefficients identified in the focal model. For a given ε', δ is solved for and used to adjust the regression coefficients. This approach ensures that both coefficients undergo equally relative adjustments to arrive at the desired elasticity.

For each ε' and δ, the model is re-estimated to project the number of average entrants using the new purse and similarity coefficients, the baseline coefficients on all other variables. Most notably, the influence of competitive expectations is held fixed; however, when re-estimating the model, competitive expectations are re-estimated as well so that scientists are implicitly all aware of the change in the elasticity and its role on entry decisions.

References

Azoulay, P., Fons-Rosen, C., and Graff Zivin, J. S. (2016). Does science advance one funeral at a time? Technical report, National Bureau of Economic Research.

Azoulay, P., Furman, J. L., Krieger, J. L., and Murray, F. (2015a). Retractions. *Review of Economics and Statistics*, 97(5):1118–1136.

Azoulay, P., Graff Zivin, J., and Wang, J. (2010). Superstar extinction. *The Quarterly Journal of Economics*, 125(2):549–589.

Azoulay, P., Graff Zivin, J. S., Li, D., and Sampat, B. N. (2015b). Public R&D investments and private-sector patenting: Evidence from NIH funding rules.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49.

Bhattacharya, J. and Packalen, M. (2011). Opportunities and benefits as determinants of the direction of scientific research. *Journal of Health Economics*, 30(4):603–615.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., and Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*.

Ding, W. W., Murray, F., and Stuart, T. E. (2006). Gender differences in patenting in the academic life sciences. *Science*, 313(5787):665–667.

Gavetti, G. and Levinthal, D. (2000). Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly*, 45(1):113–137.

Hegde, D. and Sampat, B. (2015). Can private money buy public science? disease group lobbying and federal funding for biomedical research. *Management Science*, 61(10):2281–2298.

Lin, J. and Wilbur, W. J. (2007). Pubmed related articles: a probabilistic topic-based model for content similarity. *BMC Bioinformatics*, 8(1):1.

Nelson, R. R. and Winter, S. G. (1982). *An evolutionary theory of economic change*.

Stern, S. (2004). Do scientists pay to be scientists? *Management Science*, 50(6):835–853.

Stuart, T. E. and Ding, W. W. (2006). When do scientists become entrepreneurs? the social structural antecedents of commercial activity in the academic life sciences. *American Journal of Sociology*, 112(1):97–144.