NBER WORKING PAPER SERIES

DOES THE NIH FUND EDGE SCIENCE?

Mikko Packalen
Jay Bhattacharya

Working Paper 24860
http://www.nber.org/papers/w24860

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2018

We thank Bruce Weinberg, Joel Blit, Jeremy Goldhaber-Fiebert, David Studdert, Partha Bhattacharyya, Neesha Joseph, and Walter Schaeffer for discussions. We also thank seminar participants at the Institute for Fiscal Studies, University of Illinois at Chicago Institute of Government and Public Affairs, Stanford Medical School, and the National Bureau of Economic Research working group on Invention in an Aging Society for helpful feedback. Finally, we thank the National Institute of Aging for funding for this research through grant P01-AG039347. We are solely responsible for the content and errors in the paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Does the NIH Fund Edge Science?
Mikko Packalen and Jay Bhattacharya
NBER Working Paper No. 24860
July 2018
JEL No. I15,O31

ABSTRACT

The National Institutes of Health (NIH) plays a critical role in funding scientific endeavors in biomedicine that would be difficult to finance via private sources. One important mandate of the NIH is to fund innovative science that tries out new ideas, but many have questioned the NIH's ability to fulfill this aim. We examine whether the NIH succeeds in funding work that tries out novel ideas. We find that novel science is more often NIH funded than is less innovative science but this positive result comes with several caveats. First, despite the implementation of initiatives to support edge science, the preference for funding novel science is mostly limited to work that builds on novel basic science ideas; projects that build on novel clinical ideas are not favored by the NIH over projects that build on well-established clinical knowledge. Second, NIH's general preference for funding work that builds on basic science ideas, regardless of its novelty or application area, is a large contributor to the overall positive link between novelty and NIH funding. If funding rates for work that builds on basic science ideas and work that builds on clinical ideas had been equal, NIH's funding rates for novel and traditional science would have been the same. Third, NIH's propensity to fund projects that build on the most recent advances has declined over the last several decades. Thus, in this regard NIH funding has become more conservative despite initiatives to increase funding for innovative projects.

Mikko Packalen
University of Waterloo
Department of Economics
200 University Avenue West
Waterloo, ON N2L 3G1
Canada
packalen@uwaterloo.ca

Jay Bhattacharya
117 Encina Commons
CHP/PCOR
Stanford University
Stanford, CA 94305-6019
and NBER
jay@stanford.edu

1. Introduction

The United States National Institutes of Health (NIH) both supports and sets the agenda of biomedical science. With an annual budget of over $37 billion, NIH funds the work of 300,000 researchers and scientists across the globe (National Institutes of Health, 2017a). As such, the NIH plays a pivotal role in setting the incentives that biomedical scientists have to try out novel ideas in their work, a vital aspect of fruitful scientific investigation (Collins 2015; Sampat 2012).

The regular exercise of trying out new ideas, which we term *edge science*, is a key contributor to the advance of scientific disciplines (e.g. Marshall 1929; Usher 1920; Kuhn 1962, 1977). The willingness of scientists in a field to try out new ideas generated by others is at least as important as the generation of ideas. For when new ideas are first born, they are often raw and poorly understood (Kuhn 1962, 1977). Hence, the attention of many scientists is needed to help an idea develop from the germ of an idea to a transformative idea. Extensive debate over a new idea is needed to develop it, to test its validity, and to determine the areas where the new idea might be worthwhile. Ideas that have been vetted at least partially are also more likely to be seen by fellow scientists as likely to produce important results, and thus draw the attention of other scientists.

For scientists, however, edge science is inherently risky, since it can be difficult to predict whether a new idea will produce fruitful results. Working scientists often point to failure as a precursor to success, but there is no guarantee that a particular idea will work (e.g. Livio 2013; Firestein 2015; Popovian 2016; Zaringhalam 2016). Public support for edge science, even if failure is likely, can help establish appropriate incentives for novel work in biomedical science.

Moreover, without explicit incentives to pursue edge science, there is too little novel work in science due to the coordination problem that is inherent in the formation of robust scientific communities to new areas of investigation (Besancenot and Vranceanu 2015; Packalen and Bhattacharya 2017). Thus, in effect, NIH support for the trying out of new ideas in science can solve a failure in the market for the production of novel science.

While NIH strives to "foster fundamental creative discoveries [and] innovative research strategies" (National Institutes of Health 2017b), some researchers have

questioned its ability to effectively recognize and cultivate truly groundbreaking science (e.g. Cook-Deegan 1996; Nicholson and Ioannidis 2012; Kolata 2009). Based on the types of research and researchers that the NIH funds, some have even concluded that the NIH underfunds "discovery, creativity and innovation" (Cech 2015).

Both scientific and political considerations may lead the NIH to underfund the trying out of new ideas. First, because the NIH visibly spends public money, it needs to show discrete manifestations of improvements in health, as well as technological breakthroughs, arising from its supported research (Hedge and Sampat 2015). This consideration leads to preference to support ideas that have already shown promise rather than edge science. Second, NIH scientific review panels, for reasons related to their constitution, tend to reward projects that are evidently feasible over projects that are novel (e.g. Langer 2012; Nicholson and Ioannidis 2012; Alberts et al. 2014; Geman and Geman 2016). In addition, the changing demographics of biomedical scientists can exacerbate potential gaps in support for novel work (e.g. Kaiser 2014; Daniels 2015; Lauer et al. 2017; Levitt and Levitt 2017). To the extent that early-career scientists pursue novel research paths more often than do experienced scholars, the graying of grant recipients can lead to decreased support for novel science.

In this paper, we provide a quantitative assessment of whether NIH policies encourage or impede the trying out of new ideas. We first determine the novelty of each biomedical publication based on its textual content. We then compare the frequency of NIH funding for novel papers against the frequency of NIH funding for papers that build on well-established ideas. This comparison reveals to what extent NIH funding practices reward the pursuit of edge science.

Our findings offer qualified support for NIH's stated mission to support innovative research. We find that novel contributions are more often NIH supported than are contributions that only build on well-established ideas. But while this positive link between novelty and NIH funding is present for contributions that build on novel basic science ideas, it is absent from contributions that build on novel clinical ideas. Moreover, once we account for the differences in funding rates across idea types (for example, contributions that build on advances in genomics receive NIH support much more often than do contributions that only build on clinical ideas), we find that novel science and

traditional science receive NIH support at about the same rate. This finding suggests that typically the NIH is able to direct funds to innovative areas of investigation (such as both basic and applied work on genomics) but the NIH is not able to direct funds to innovative individual projects within fields. In addition to these substantive findings, our analysis contributes by putting forward a systematic approach that can be used to quantify the novelty preferences of any funding agency.

Despite the importance of NIH support of edge science, there has been surprisingly little research quantifying its extent. Prior research has focused on measuring the effect of NIH funding on the productivity of scientists as measured by publications, citation counts, and patents (e.g. Murray et al. 2016; Jacob and Lefgren 2011a, 2011b; Azoulay et al. 2011, 2017; Blume-Kohout 2012; Blume-Kohout et al. 2015; Li 2017; Li and Agha 2015; Li et al. 2017). Though these outputs are important, they are not quantifiable measures of edge science.

The focus in citation analyses of NIH funding, for instance, has been on capturing the production of high-impact science rather than what kind of science is being produced – novel or more conventional. While work on identifying NIH support for such acts of genius is important, from the perspective of optimal science policy it is *also* valuable to understand whether NIH succeeds in its aim of favoring novel science. For, as argued above, advances that later turn out to be transformative are often not well elaborated when they are first born. Scientists who are willing to try out the new ideas as inputs to their research thus play an important role in developing the ideas to mature advances. As argued above, the NIH can serve a useful role by funding such edge science. It is thus valuable to offer a quantitative assessment of NIH's novelty preferences.

Our approach to identifying edge science follows Packalen (2018), which examines how the tendency to pursue edge science varies across countries. This approach allows each paper included in the analysis to contribute on multiple dimensions of knowledge, and the novelty of an idea input is determined relative to application of other comparable ideas in the same research area. This methodology extended the approach developed in Packalen and Bhattacharya (2016, 2017).

The balance of the paper proceeds in the standard order: methods, results, discussion, and conclusion.

2. Methods

Our analysis covers over 24 million research articles in the MEDLINE database. MEDLINE is a comprehensive database of published biomedical research articles published between 1950-2017. To determine the ideas upon which each paper builds, we employ the Unified Medical Language System ("UMLS") metathesaurus, which is a curated and controlled vocabulary of over 5 million biomedical terms. For each paper in the MEDLINE database, we determine which UMLS terms appear and what idea type those terms represent. We determine the idea type based on the UMLS semantic category of each term. There are 127 UMLS semantic categories for terms; among these categories are "Gene and Genome", "Neoplastic Process", and "Quantitative Concept.". In our analysis, each UMLS semantic category represents one idea type. We determine the research area of the paper based on the journal where the paper was published. Journal category is assigned by the National Library of Medicine. There are 125 categories, including "Cardiology", "General Surgery" and "Molecular Biology."

We conduct our analysis at a contribution level, where *contribution* is defined as a link from a paper to an (idea type, research area) pair. The underlying assumption is that each paper that is linked to an (idea type, research area) pair contributes to the scientific debate of how the idea functions when applied to the specific research area. When a paper is linked to multiple (idea type, research area) pairs, we assume that the paper represents multiple contributions because the paper has the potential to contribute to multiple research areas and idea types – it can enhance our understanding on multiple dimensions.

To determine which contributions represent novel science, we determine the vintage of each term based on the year that the term first appears in the MEDLINE database. We refer to the year of first appearance of each idea as the *cohort* of the idea. For each contribution we then determine the cohort of the *newest idea input* that is linked to the contribution, which we refer to as *idea input vintage*. In our approach, we consider contributions that build on at least one relatively recent idea to be novel, while contributions that build on well-established ideas represent more traditional science.

This methodology for identifying novel contributions from the text of research articles builds on several earlier analyses. The closest related analysis is Packalen (2018), which introduced the contribution-level approach for identifying edge science. We refer the reader to this earlier work for lists of examples of ideas captured by this approach as well as descriptive statistics on the MEDLINE database in general and on the estimated vintage of these ideas and on how often the ideas have been used in the biomedical literature. A related previous analysis had employed a paper-level approach for identifying novel ideas using the UMLS metathesaurus and the MEDLINE database (Packalen and Bhattacharya 2017), which in turn built on Packalen and Bhattacharya (2016) in which an n-gram approach was used to identify ideas and novel contributions based on the text of articles in the MEDLINE database. This methodology identifies a nearly complete history of the ideas of biomedicine (both important and obscure), going back to the middle of the 20th century.

We identify the NIH funding status of each paper from the MEDLINE database. A paper where any of its authors is reported to have acknowledged NIH funding for the project is assigned "with NIH funding" status, and any paper where none of the authors is reported to have acknowledged NIH funding for the project is assigned "without NIH funding" status. The advantage of using paper-level NIH funding acknowledgements as the source of NIH funding information, instead of data on funded NIH grant proposals, is that the former gives information on projects that were actually completed with NIH funding. This distinction matters because of a potential disconnect between funding proposals and actual research. If a scientist agrees with the commonly held perception that NIH review committees are conservative in funding decisions, but is inclined to pursue novel work nevertheless, it may make sense to hide the novelty – at least to some degree – in funding applications. In any case, the type of work scientists do is more important – as far as the social benefits of science are concerned – than the work that scientists promise to do when applying for funding. Grantmanship is an important topic but is largely separate from the focus of the analysis pursued here.

We limit the analysis to papers which first author has a U.S. affiliation (for papers published before 2014 MEDLINE mostly has affiliation information for the first author). This restriction is important because scientists working in the U.S. are much more likely

to build on new ideas compared to the average across all countries in biomedicine (Packalen 2018) and because the NIH disproportionately funds U.S. biomedical scientists. Hence, if non-U.S. papers were included in the analysis, the results would reflect also the relative scientific frontier position of U.S. scientists on average and the NIH's differential propensity to fund U.S. vs. non-U.S. scientists. By focusing only on papers which first author is in the U.S., our results will reflect a comparison of the vintage of ideas that NIH funded scientists tend to build on in their work (i.e. their propensity to pursue edge science) and the vintage of the ideas that comparable scientists without NIH funding tend to build on in their work.

3. Results

3.1 Aggregate Relationship Between Novelty and NIH Funding

We first examine how the share of NIH funded contributions varies by the idea input vintage. Recall that idea input vintage captures the year that the newest idea that the contribution builds upon was introduced in the literature. Our goal is to uncover whether NIH funding is directed disproportionately to projects that try out ideas introduced to the literature more recently, in line with the NIH's goal of supporting innovative science. The overall results for contributions across all idea types and research areas are shown in Figure 1 for contributions published between 2010 and 2016. The figure depicts the fraction of contributions that are in papers with NIH funding (the vertical axis) as a function of the idea input vintage (the horizontal axis). The horizontal dashed line in Figure 1 represents the average share of contributions funded by the NIH over all cohorts.

The results in Figure 1 suggest an inverted "U" shape to the relationship between NIH funding and novelty of idea inputs. The share of research supported by the NIH is the highest for papers that build on ideas that are relatively new but not too recent; there appears to be a substantial time lag for funding work on a new idea. The NIH funding rate is lower for contributions that build on the most recent ideas (post-2005 cohorts) than it is for contributions that build on a bit more mature ideas (1990-2005 cohorts), but the share of NIH support is lowest for contributions that only reference well-established

knowledge i.e. ideas introduced to the literature a long time ago (pre-1970 cohorts). Thus, NIH funding practices seem to favor novel work with a 7 to 10-year delay. The magnitude of these idea-vintage related NIH funding differences is also considerable: the NIH funding rate is over 55% for contributions that make use of 10 to 25-year-old ideas, whereas it is only about 45% for contributions that either build on only well-established knowledge or build on some very recent ideas.

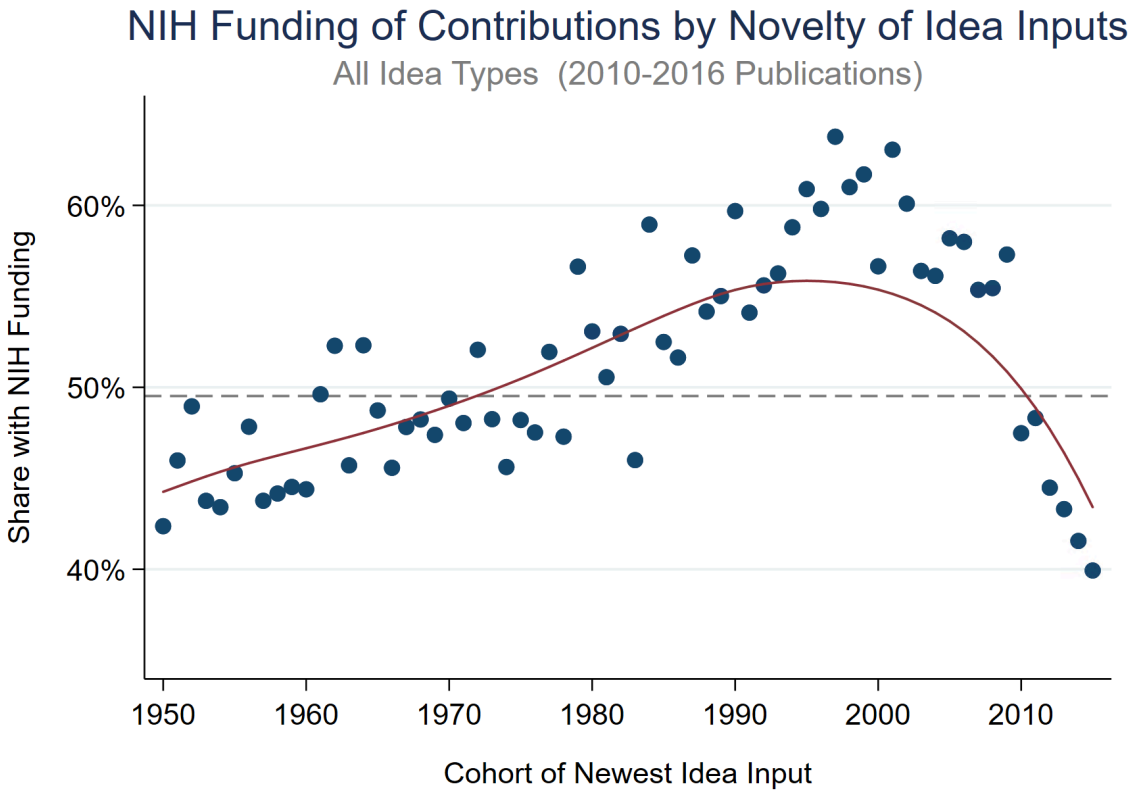

Figure 1: Share of NIH Funding by Novelty of Idea Inputs (2010-2016): All Idea Types. Calculated based on 1,478,475 biomedical research papers published during 2010-2016. The horizontal axis captures the idea input vintage (the cohort of a contribution is the year when the newest idea input used in the contribution was introduced to the literature). Later (earlier) cohort years represent more novel (more conventional) science. The vertical axis captures the rate of NIH funding. The markers capture the mean NIH funding rate for each idea input vintage. The solid line represents a non-parametric regression line estimate. The dashed line represents the average funding rate across all cohorts.

Figure 2 shows the same relationship for two earlier decades, namely 1990s (top left panel) and 2000s (top right panel), as well as for two sub-periods of the current decade,

namely years 2010-2014 (bottom left panel) and 2015-2016 (bottom right panel). The results in Figure 2 show that also during the earlier decades the NIH funded novel science at higher rates than it funded traditional science. However, unlike in the 2010s, in the 1990s the NIH funded even contributions that built on the most recent ideas at a higher rate than it funded contributions that only built on well-established ideas. That is, the funding penalty for the most novel work appears to be a recent phenomenon – in the 1990s there was no delay in funding brand new ideas. Thus, in this regard the NIH appears to have become more conservative in its support for edge science over time.

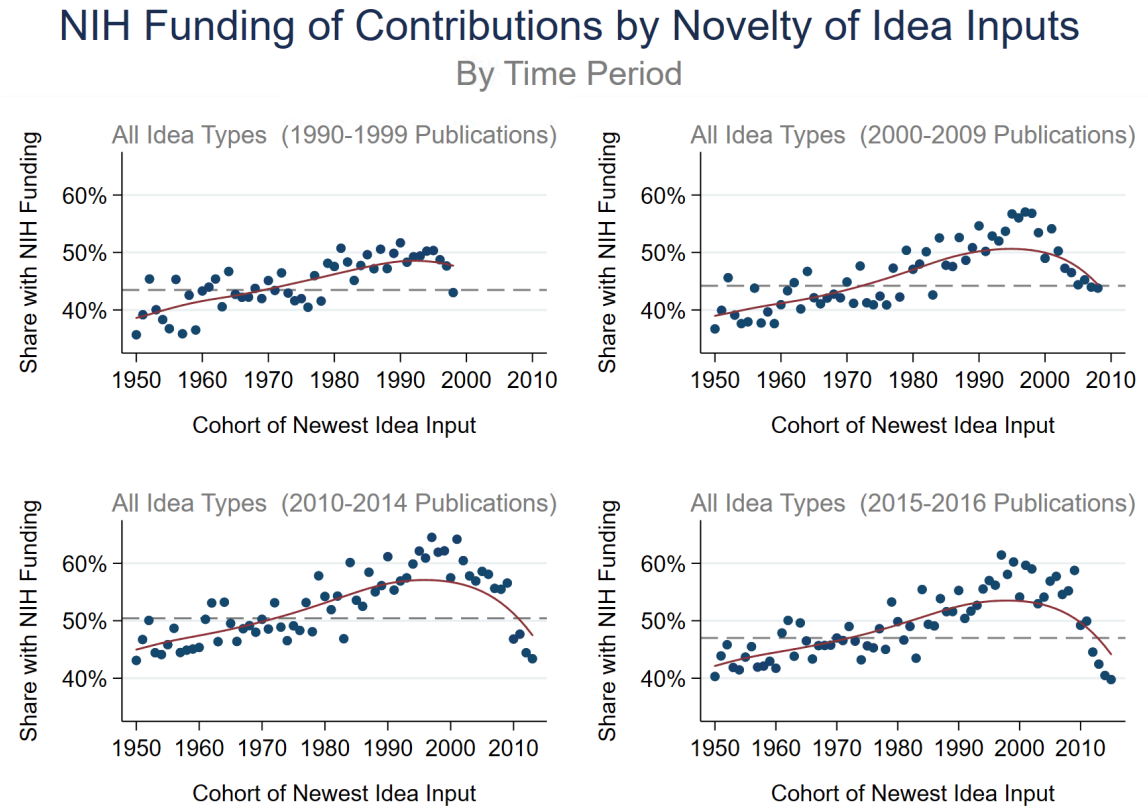

Figure 2: Share of NIH Funding by Novelty of Idea Inputs: for 1990s (top left), 2000s (top left), 2010-2014 (bottom left) and 2015-2016 (bottom right). Results in these panels are calculated based on 1,651,429 (1,049,553) [1,093,170] {385,305} biomedical research papers published during 1990-1999 (2000-2009) [2014-2015] {2015-2016}. In each panel, the horizontal axis captures the idea input vintage, with later (earlier) cohort years representing more novel (more conventional) science, and the vertical axis captures the rate of NIH funding. The markers capture the mean NIH funding rate for each idea input vintage. The solid line represents a non-parametric regression line estimate. The dashed line represents the average funding rate across all cohorts.

3.2 Link Between Novelty and NIH Funding by Idea Type and Research Area

Figure 3 shows the relationship between novelty and NIH funding separately for four idea types: Genes and Genome (top left panel), Amino Acid, Peptide, or Proteins (top right panel), Neoplastic Process (bottom right panel), and Pharmacologic Substance (bottom left panel). Like in Figure 1, these results are for papers published between 2010 and 2016. The dashed line in each panel indicates that average level of NIH support for contributions that build on an idea from the idea type.

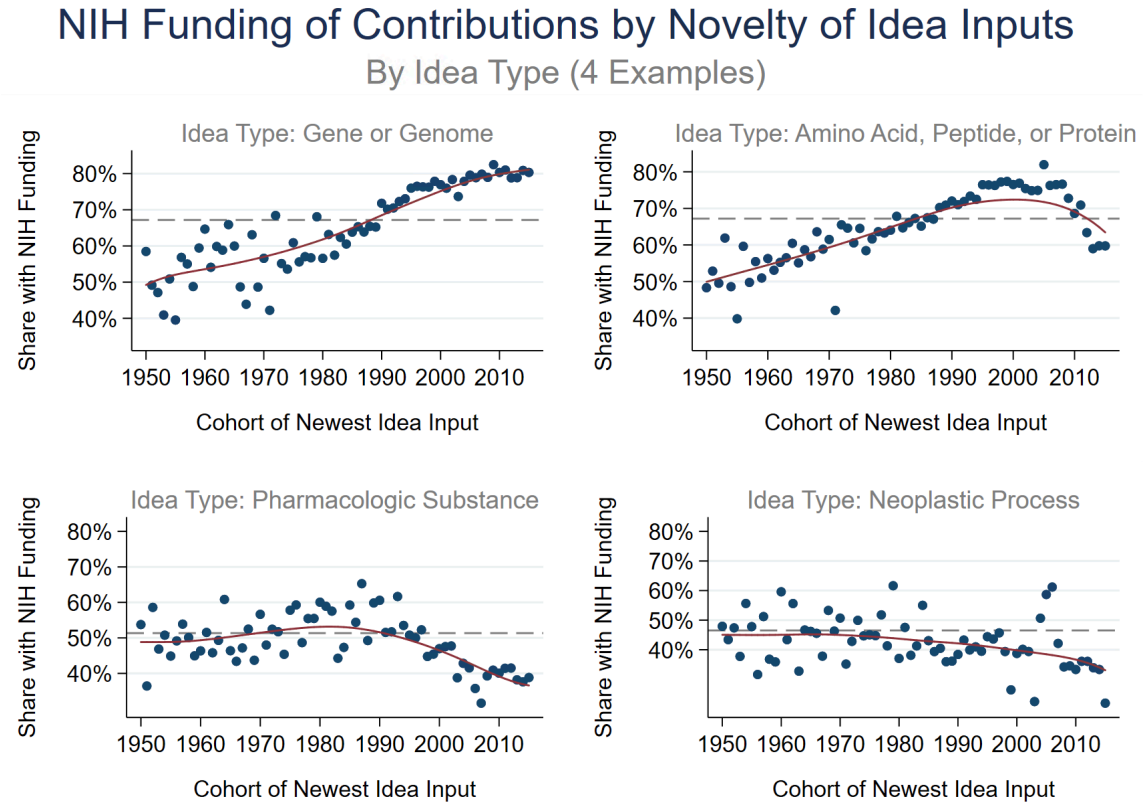

Figure 3: Share of NIH Funding by Novelty of Idea Inputs for Four Examples of Idea Types: Genes and Genome (top left), Amino Acid, Peptide, or Proteins (top right), Neoplastic Process (bottom right), and Pharmacologic Substance (bottom left). Calculated based on biomedical research papers published during 2010-2016. In each panel, the horizontal axis captures the idea input vintage, with later (earlier) cohort years representing more novel (more conventional) science, and the vertical axis captures the rate of NIH funding. The markers capture the mean NIH funding rate for every idea input vintage. The solid line represents a non-parametric regression line estimate. The dashed line represents the average funding rate across all cohorts.

The results in Figure 3 show that NIH funding practices reward novelty for work on some idea types but not for other idea types. The top panels of Figure 3 show that work that references genes or proteins is more likely to receive funding when the it mentions a relatively new gene or protein compared to work that only mentions genes or proteins first introduced to the literature a long time ago. Bottom right panel of Figure 3 shows that papers that reference ideas in the neoplastic process category are somewhat less likely to receive funding when they build on an idea in this category that is of a more recent vintage. Bottom left panel of Figure 3 shows a similar finding for drugs: papers that mention a new drug are much less likely to be NIH funded than are papers that only mention older drugs. This last finding likely reflects a decision by the NIH to focus on funding research on drugs that are no longer under patent as pharmaceutical companies have a reduced incentive to investigate the properties of such drugs.

The results in Figure 3 thus indicate that NIH overall preference for novel work is heterogeneous across idea types. This, together with the variance in average funding rates across idea types (as indicated by dashed lines in each panel of Figure 3), raises the prospect that the link between novelty and NIH funding (as shown in Figure 1) is driven by NIH's preference for funding certain idea types and research areas where a lot of progress is taking place, rather than by NIH's preference for funding novel work within idea types and research areas. Perhaps, for example, the results are driven by the fact that NIH is very favorable to any research that makes use of either novel or well-established advances in genomics (this fact is evident by the high average funding rate indicated by the dashed line in Panel A of Figure 3).

To investigate this possibility, we calculated the NIH funding rate by cohort of the newest idea while holding the average funding rate the same across all (idea type, research area) pairs. In this calculation, we allow the funding rate to vary *only across cohorts* within each (idea type, research area) pair. To achieve this objective, we first determine for each (idea type, research area) pair how often papers linked to it receive NIH support across all cohorts. We then determine, for each cohort, how much NIH funding papers linked to that cohort receive relative to the average among all cohorts linked to the same (idea type, research area) pair. We use this ratio to adjust the number of observed NIH funded papers accordingly for each (idea type, research area) pair. This

way, an NIH funded contribution linked to an (idea type, research area) pair where contributions receive funding, say, 80% of the time across all cohorts is only counted as half a contribution relative to a contribution linked to an (idea type, research area) pair where contributions receive funding 40% of the time across all cohorts. As a result, the funding rate for an (idea type, research area) pair is above the average funding rate only for those cohorts for which the funding rate of the cohort is above the average funding rate for that (idea type, research area) pair across all cohorts. The results for this calculation, which in effect controls for differences in NIH funding rates across idea types and research areas, are shown in Figure 4.

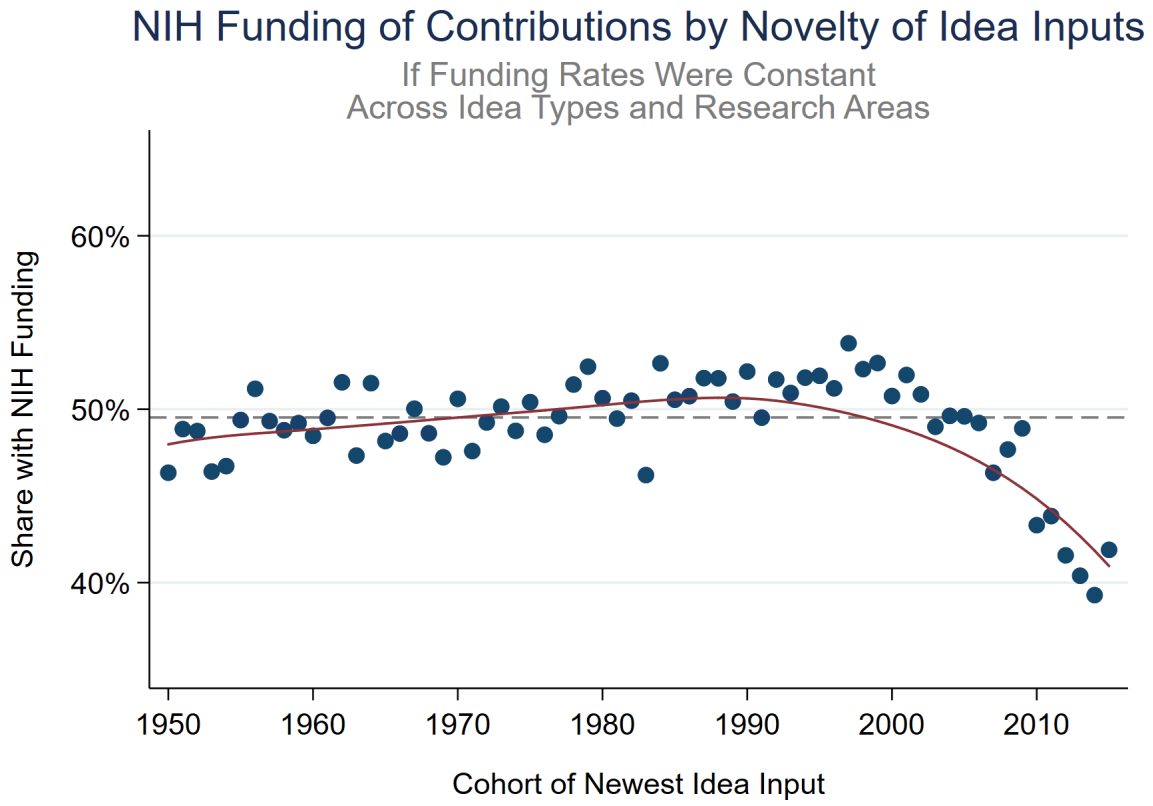

Figure 4: Share of NIH Funding by Novelty of Idea Inputs: If Funding Rates Were Constant Across Idea Types and Research Areas. Calculated based on 1,478,475 biomedical research papers published during 2010-2016. The horizontal axis captures the idea input vintage, with later (earlier) cohort years representing more novel (more conventional) science. The vertical axis captures the rate of NIH funding. The markers capture the mean NIH funding rate for each idea input vintage. The solid line represents a non-parametric regression line estimate. The dashed line represents the average funding rate across all cohorts.

The results shown in Figure 4 still indicate an inverted U-shaped relationship between novelty of idea inputs and NIH funding, but quantitatively the results are now quite different. The highest NIH funding rate for novel work is now only a little above the funding rate for contributions that build on well-established ideas, and the funding rate for work that builds on the most recent ideas is now significantly below the average funding rate across all cohorts. Together with the unadjusted results shown in Figure 1, this result suggests that while the NIH has been successful in funding innovative science, this success has been due to its differential support for certain idea types where many ideas are relatively novel (such as genomics). Conversely, the results imply that within research areas and idea types the NIH has not been able to funnel resources to the most innovative projects.

Table 1 summarizes the results for 28 largest categories of the 127 idea types represented by the UMLS semantic type taxonomy (by contrast with Figure 3 which showed idea type specific results for four idea types). The average of the results for the remaining 99 idea types are summarized in the last row. Columns 1a and 1b, respectively, list the idea type and whether the idea type represents basic science ideas, clinical ideas or miscellaneous ideas. Columns 2a-2d report descriptive statistics: the number of contributions linked to the idea type regardless of the cohort the contribution (column 2a), the share of contributions with NIH funding (column 2b), the number of novel contributions linked to the idea type (column 2c), and the share of novel contributions with NIH funding (column 2d). For these calculations, a novel contribution is defined as a contribution that builds on an idea that is at most 15 years old at the beginning of the sample period (idea cohorts 1995-2016 for the 2010-2016 sample period). The numbers across the rows in column 2b indicate that there is a lot of variation in NIH's propensity to fund ideas depending on the idea type, and that this propensity is generally higher for basic science idea types than it is for clinical idea types. The heterogeneity in NIH support across idea types also implies that our main results cannot be simply explained by the mechanical delay induced by the grant review process.

Column 3 reports the main result: a summary statistic that captures NIH's relative propensity to fund novel work. We refer to this statistic as the *edge funding ratio*. We calculate the edge funding ratio by dividing the share of novel contributions with funding

by the share of non-novel contributions with NIH funding. Here a novel contribution is defined as discussed above (i.e. a contribution that builds on an idea that is at most 15 years old), and a non-novel contribution is defined as a contribution that builds only on ideas introduced to the literature in the year 1970 or earlier. An edge funding ratio of 1.10 (0.90) means that the NIH's propensity to fund novel science is 10% higher (10% lower) than is the NIH's propensity to fund traditional science. We color each cell based on the relative propensity: red indicates an edge funding ratio above 1.20 (NIH's funding practices strongly favor novel work), orange indicates an edge funding ratio between 1.10 and 1.20 (NIH's funding practices are favorable to novel work), and blue indicates an edge funding ratio below 0.90 (NIH's funding practices favor traditional science).

The results in column 3 show that among work linked to basic science idea types, the NIH generally funds work that builds on novel ideas considerably more often than it funds traditional science. This link between novel work and NIH funding is not present for clinical idea types (with the exception of the Laboratory Procedure category) but the link appears to be present for some miscellaneous idea types including the category Quantitative Concept. These findings reinforce the findings drawn earlier based on Figure 3. The NIH's propensity to fund innovative science varies considerably across idea types, and while its propensity to fund novel science is robust for basic science idea types (most notably genetics and proteomics) it is not present for most clinical idea types.

Column 4a reports adjusted estimates from a specification which controls for differences in funding rates across idea types. The results correspond to the hypothetical scenario that NIH funding rates were equal across all idea types but may vary across research areas and across idea cohorts. The cells for individual idea types are empty because the results are the same as in the baseline specification (column 3) when only one idea type is included in the analysis. Column 4b reports estimates from a specification which controls for differences in funding rates across research areas (identified from journal categories). The results correspond to the hypothetical case that NIH funding rates were equal across research areas but might vary across idea types and across idea cohorts. Column 4c reports estimates from a specification which controls for differences in funding rates across idea types and research areas, similar to the approach employed in the analysis reported in Figure 4. The results correspond to the hypothetical

scenario that funding rates are equal for all (idea type, research area) pairs but may vary across idea cohorts within each (idea type, research area pair).

Together with the result in the first row of column 3, the results in the first row of columns 4a, 4b and 4c imply that the overall positive link between novelty and NIH funding rates is largely driven by funding difference across idea types, rather than funding differences across research areas or funding differences within specific idea types or research areas. For controlling for only differences across idea types decreases the edge funding ratio from 1.26 to 1.05, controlling for only differences across idea types decreases the edge funding ratio from 1.26 to 1.14, and controlling on both dimensions decreases the edge funding ratio to 1.04. Controlling for differences across idea types thus has the largest impact. Quantitatively, the fact that the edge funding ratio is only 1.04 after controlling for differences in funding rates across idea types and research areas implies that, on average in biomedicine the NIH's ability to direct resources to innovative projects within idea types and research areas is at best only modest. A comparison of the results in the other rows of column 3 against the results in the corresponding row of column 4b further demonstrates that controlling for funding rate differences across research areas has only a modest impact on the edge funding ratio.

Columns 5a-5d report results for four time periods: 1990s, 2000s, 2010-2014, and 2015-2016. For most idea types, the results are relatively stable over time, indicating that variation across idea types reflects systematic differences in NIH's propensity to fund novel work rather than random variations.

Table 2 reports the edge funding ratio for different (idea type, research area) pairs, with idea type on the vertical axis and research area on the horizontal axis. By contrast, Table 1 showed edge funding ratios by idea type. The columns of Table 2 report the edge funding ratio for each of the 22 largest research areas, as determined by the number of contributions linked to each journal category; the last column summarizes the result for the 103 remaining research areas. The cells are left empty for those (idea type, research area) combinations for which there are less than 100 novel contributions linked to it. The results in the first and last row, as well as the rows for the three idea types with the most novel contributions linked to them (genes, proteins, and drugs) show little variation across research areas.

This result reinforces the earlier finding that most of the variation in the novelty-NIH funding link is driven by differences across idea types rather than by differences across research areas. For other idea types, there is a bit more variation across research areas (and also many more empty cells) but this likely mostly reflects random variations due to the small number of novel contributions based on which the estimates for many such cells are calculated rather than systematic variations in the link between novelty and NIH funding across research areas.

4. Discussion

Taken together, our results show that, while the NIH continues to provide support for work on new ideas in biomedicine, its support for the most novel ideas has waned in recent decades. Furthermore, much of NIH's support for novelty stems from field-specific funding priorities, with fields that are more likely to have featured fast progress in recent decades (such as genomics) receiving greater NIH support compared to fields that have not experienced as fast a rate of scientific progress in recent decades. Less positively, our results show that within fields, the NIH generally does not discriminate in its support between newer and older ideas.

Our results provide indirect support for the idea that the NIH review committees are risk-averse: within given fields, biomedical research contributions that build on brand new ideas report having NIH support less often than do contributions that build on ideas that are relatively new but have had some opportunity to mature. A plausible explanation for this pattern is that contributions in the former category – the most novel projects – are the most likely to be seen as infeasible or too risky.

However, our evidence also suggests that there is considerably heterogeneity in the NIH's attitude toward novelty depending on the type of the idea that a project builds upon. Projects that build on a novel basic science idea (especially a recently discovered gene or protein) are more likely to be funded than are projects that build on comparable older ideas. By contrast, the novelty-NIH funding link is not present for clinical ideas. Thus, while we find consistent NIH support for edge science linked to genomics and concepts in subcellular biology more generally, other types of edge science in

biomedicine are generally less likely to be NIH supported than are comparable projects that build on more traditional ideas.

Our analysis does not reveal the mechanisms behind this heterogeneity. It is well known that NIH has provided generous support for research linked to genomics in recent decades (e.g. Joyner et al. 2016). But this preference to fund genomics related work does not explain the positive link between novelty and frequency of NIH support among projects that build on advances in genomics. Hence, further research on the mechanisms that determine the presence of a link between NIH support and edge science is warranted.

The fact that, after controlling variations in in funding rates across idea types and research areas, there is no link between novelty and NIH funding suggests, taken as a whole, NIH review panels are either unwilling to reward novelty or are not able to discern which research projects are more novel. This is true despite the fact that innovation is one of the main criteria that review panels are asked to use in evaluating grant applications. To the extent that rewarding novelty within idea types and research areas is infeasible, it may be that the only feasible approach for directing funds to novel work is to provide ample funding to entire scientific fields that are thought to be advancing at a faster rate than others. For example, current funding rates for genomics related work – which are significantly higher than funding rates are in general – could be justified by the fact that in recent decades this area has advanced at a faster rate than have other areas.

There are, of course, inherent risks in directing a disproportionate share of funds to certain scientific fields. Many of these risks have been discussed in the context of examining the desirability of projects that exemplify "big science" and "big biology" (e.g. Rosbash 2011; Alberts 2012; Petsko 2009; Vermeulen et al. 2010; Peifer 2017; Joyner et al. 2016). When a top-down approach to research resource allocation is adopted, the funds may end up being directed to areas that turn out to be less fruitful than expected, at the expense of funding a more diverse and thus less risky portfolio of research projects. Moreover, inertia in funding decisions will likely keep some areas well-funded long past the eventual stagnation of these fields. It is also questionable whether funding agencies or anyone can know in advance which discipline or organism holds the key to the next important development in biomedicine.

These risks of a less diversified funding portfolio must be weighed against the benefits. Our finding that the novelty-NIH funding link is at best fragile when comparing contributions within idea types and research areas suggest that the alternative to top-down research resource allocation – letting individual scientists choose research directions more freely and trying to direct funds to innovative projects rather than innovative fields – may not be any better an approach in terms of facilitating vibrant scientific communities that have strong incentives to pursue also novel research paths.

There are at least two potential explanations for our finding that over the last three decades NIH has become more conservative in its support for work that builds on the very latest ideas. First, it may be that review committees have become more cautious in terms of funding very novel work. This is consistent with related recent results that have indicated that biomedical researchers themselves have become more conservative in their research choices (Rzhetzky et al. 2015). Second, perhaps the newest ideas that scientists have generated in the 2000s have not been as widely and rapidly accepted by the scientific community as the newest ideas that were generated during 1990s. This second explanation is consistent with concerns about stagnation in technological progress during the 2000s in general, and in medicine more specifically (e.g. Le Fanu 2011), a hypothesis that is supported by empirical evidence showing that health advances have become less common and harder to achieve in recent decades (Bloom et al. 2017).

We cannot tell which of these explanations is correct (or if either are correct) from the results we present in the paper, though our results contribute to this discussion. However, if researchers and funding agencies have become more conservative at the same time as biomedical research has stagnated, it is that much more important to identify ways to incentivize researchers and funding agencies to pursue and support novel work that builds on new ideas. The approach employed in this paper is useful for this purpose as it demonstrates how to identify innovative research based on research texts and how to use that as a basis for evaluating which funding agencies and funding practices reward innovative research projects.

5. Conclusion

The trying out of new ideas is a fundamental feature of healthy scientific disciplines; scientific progress depends upon a scientific community being open to trying out new ideas continuously and systematically. Absent the attention of many scientists, promising new ideas cannot be effectively evaluated, developed, and tested for unexpected applications. Yet, for an individual scientist the decision to try out a new idea is risky. Work that builds on well-established ideas is often the safer choice, as the results the latter be more predictable and materialize sooner, even as innovative work that builds on newer ideas may ultimately prove to have a higher impact.

The United States National Institutes of Health (NIH) has played a pivotal role in advancing public health and medical research in the past several decades. As the world's largest source of medical research funding, it has supported the development of numerous groundbreaking medical technologies and pharmaceuticals that are in common use today.

One of the primary goals of NIH funding is to support innovative biomedical research, including work that tries out new ideas, which we term edge science. But because the NIH faces pressure to deliver short-term successes and because identifying successful novel projects may be more difficult, many have questioned the NIH's ability to fund innovative science. At the same time, quantitative analyses of whether NIH actually achieves this important mission have been scarce in the published literature.

Our analysis of published biomedical research articles in the comprehensive MEDLINE database finds strong support for the preference that the NIH has for innovative work. Those contributions that build on relatively recent ideas are substantially more likely to have been funded by the NIH compared to contributions that only rely on older, well-established, ideas. On average, the NIH thus succeeds in providing support for edge science.

Despite this rosy evaluation, our findings suggest that there is room for the NIH to do more to promote innovative science. We find that in the current funding environment, there is a time lag for NIH funding of papers that rely on new ideas: work that builds on brand new ideas is funded at a lower rate than is work that builds on ideas that have had a chance to mature at least 5-10 years. One reason for this may be that NIH review

committees are risk averse; they are less likely to support projects that rely on brand new ideas (which may be the riskiest) than ideas that have had some opportunity to mature (which properties have thus become a bit better understood).

This pattern has not always been true of NIH funded papers. In the 1990s, papers that relied on brand new ideas were the most likely to have received NIH funding. NIH review committees and funding decisions have thus have become more conservative and risk-averse over time, despite a variety of policies that the NIH has implemented in the past two decades to reward innovative and high-risk project proposals (e.g. Larkin 2003; Steinbrook 2009; Avorn and Kesselheim 2011; Woodward 2011).

We also find that NIH's ability to fund innovative science is driven mostly due to its decision to fund rapidly advancing areas of investigation such as genomics and its applications. By contrast, our findings imply that the NIH's ability to direct funding to novel projects over comparable more traditional projects within the same area of investigation is quite limited. An important implication of this finding is that a move to a more egalitarian funding model – one that doesn't favor some areas of investigation as much as the current model does – would likely lead to a significant decrease in novelty of NIH supported research, unless such a funding policy change is accompanied by a drastic shift in funding review procedures. The shift in review policies would have to render the process much more favorable to novel work within each area of investigation. Hence, the decision regarding the level of funding to each area of investigation should not be separated from the decision regarding how to arrange the review process within each area of investigation.

Our text-based approach for identifying innovative research suggests a fruitful agenda for the quantitative analysis of funding agencies. Constructively, the quantitative measurement of new idea adoption can help the NIH better understand the impact that its review practices and funding decisions have on the scientific enterprise. The methodology that we introduce in this paper can help design policies to counteract the long-term trend that we identify in the funding of work that builds on brand new ideas, and thus potentially speed scientific advance. Our methods could also be applied to NIH grant application data – especially applications that have just missed the funding threshold – to gain insights into what an increase in the NIH budget would yield in terms

of increased pursuit of innovative science. While we have focused on work that is novel in the sense that it builds on new ideas, the analysis could be extended to capture also work that tries out novel combinations of old ideas. Finally, the methodology we have put forward in this paper for evaluating the novelty preferences of a funding agency can also be applied to agencies such as the National Science Foundation that operate on scientific fields outside of biomedicine.

References

Alberts, B. (2012) "The End of Small Science?" *Science* 338: 171.

Alberts, B., Kirschner, M.W., Tilghman, S. and H. Varmus (2014) "Rescuing US Biomedical Research from Its Systemic Flaws," *Proceedings of the National Academy of Sciences* 111(16): 5773-5777.

Avorn, K. and A. S. Kesselheim (2011) "The NIH translational research center might trade public risk for private reward," *Nature Medicine* 17(10): 1176.

Azoulay, P., Graff Zivin, J. S., Li, D. and B. N. Sampat (2017) "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules," NBER Working Paper No. 20889.

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

Besancenot, Damien, and Vranceanu, Radu (2015) "Fear of Novelty: A Model of Strategic Discovery with Strategic Uncertainty," *Economic Inquiry* 53(2): 1132-1139.

Blume-Kohout, M. (2012) "Does Targeted, Disease Specific Public Research Funding Influence Pharmaceutical Innovation?" *Journal of Policy Analysis and Management* 31(3): 641-660.

Blume-Kohout, M., Kumar, K. B. and N. Sood (2015) "University R&D Funding Strategies in a Changing Federal Funding Environment," *Science and Public Policy* 42: 355-368.

Cech, T. R. (2005) "Fostering Innovation and Discovery in Biomedical Research," *JAMA* 294(11): 1390-1393.

Collins, F. R. (2015) "Exceptional Opportunities in Medical Science: A View from the National Institutes of Health," *JAMA* 313(2): 131–132.

Cook-Deegan, R. M. (1996) "Does NIH need a DARPA?" *Issues in Science and Technology* 13(2): 25-28.

Daniels, R. J. (2015) "A Generation at Risk: Young Investigators and the Future of the Biomedical Workforce," *Proceedings of the National Academy of Sciences* 112(2): 313-318.

Firestein, S. (2015) *Failure: Why Science is So Successful*. Oxford University Press.

Geman, D. and S. Geman (2016) "Science in the Age of Selfies," *Proceedings of the National Academy of Sciences* 113(34): 9384-9387.

Hegde, D. and B. N. Sampat (2015) "Can Private Money Buy Public Science? Disease Group Lobbying and Federal Funding for Biomedical Research," *Management Science* 61(10): 2281-2298.

Jacob, B. A. and L. Lefgren (2011a) "The impact of NIH postdoctoral training grants on scientific productivity," *Research Policy* 40(6): 864-874.

Jacob, B. A. and L. Lefgren (2011b) "The Impact of Research Grant Funding on Scientific Productivity," *Journal of Public Economics* 95(9-10): 1168-1177.

Joyner, M. J., Paneth, N. and J. P. A. Ioannidis (2016) "What Happens When Underperforming Big Ideas Become Entrenced?" *JAMA*, 316(13): 1355-1356.

Kaiser, J. (2014) "A Call for NIH Youth Movement," *Science* 346(6206): 150.

Bloom, N., Jones, C. I., Van Reenen, J. and M. Webb (2018) "Are Ideas Getting Harder to Find?" Manuscript.

Kolata, G. (2009) "Grant System Leads Cancer Researchers to Play It Safe," *New York Times*, 28 June.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*. Chicago University Press, Chicago.

Kuhn, T. S. (1977) Objectivity, Value Judgment and Theory Choice; in Thomas S. Kuhn, ed., *The Essential Tension*, University of Chicago Press, Chicago, pp. 320-339.

Langer, J. S. (2012) "Enabling Scientific Innovation," *Science* 338: 171.

Larkin, M. (2003) "Can Zerhouni create a bold, risk-taking NIH?" *Lancet* 362: 1382–1383.

Lauer, M., Tabak, L. and F. Collins (2017) "The Next Generation Researchers Initiative at NIH," *Proceedings of the National Academy of Sciences* 114(45): 11801-11803.

Le Fanu, J. (2010) "Science's Dead End," *Prospect*, August 10.

Levitt, M. and J. M. Levitt (2017) "Future of Fundamental Biomedical Discovery in US Biomedical Research," *Proceedings of the National Academy of Sciences* 114(25): 6498-6503.

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

Li, D. and L. Agha (2015) "Big Names or Big Ideas? Do Peer-Review Panels Select the Best Science Proposals?" *Science* 348(6233): 343-348.

Livio, M. (2013) *Brilliant Blunders: From Darwin to Einstein – Colossal Mistakes by Great Scientists that Changed Our Understanding of Life and Universe.* Simon & Schuster.

Marshall, A. (1920) *Principles of Economics*. Macmillan and Co., London.

Murray, M., Aghion, P., Dewatripont, M. Koev, J. and S. Stern (2016) "Of Mice and Academics: Examining the Effect of Openness on Innovation" *American Economic Journal: Economic Policy* 8(1): 212-52.

National Institutes of Health (2017a) "Budget" (https://www.nih.gov/about-nih/what-we-do/budget; accessed December 5, 2017).

National Institutes of Health (2017b) "Mission and Goals" (https://www.nih.gov/about-nih/what-we-do/mission-goals; accessed 5, December 2017).

Nicholson, J. M. and J. P. A. Ioannidis (2012) "Conform and Be Funded," *Science* 492: 34-36.

Packalen, M. (2018) "Edge Factors: Scientific Frontier Positions of Nations," Manuscript.

Packalen, M. and J. Bhattacharya (2016) "Age and the Trying Out of New Ideas," *Journal of Human Capital*, forthcoming.

Packalen, M. and J. Bhattacharya (2017) "Neophilia Ranking of Scientific Journals," *Scientometrics* 110: 43-64.

Petsko, G. (2009) "Big Science, Little Science," *EMBO Reports* 10(12).

Peifer, M. (2017) "The Argument for Diversifying NIH Grant Portfolio," *Molecular Biology of the Cell* 28: 2935-2940.

Popovian, R. (2016) "Dedicated Scientists Driven to Discover Cures" *Morning Consult* (https://morningconsult.com/opinions/dedicated-scientists-driven-to-discover-cures/?source=acsh.org; accessed December 5, 2017).

Rosbash, M. (2011) "A Threat to Medical Innovation," *Science* 333: 136.

Rzhetzky, A., Foster, J. G., Foster, I. T., and J. A. Evans (2015) "Choosing experiments to accelerate collective discovery," *Proceedings of the National Academy of Sciences* 112: 14569–14574.

Sampat, B. N. (2012) "Mission-Oriented Biomedical Research at the NIH" *Research Policy* 41(10): 1729-1741.

Steinbrook, R. (2009) "Opportunities and Challenges for the NIH -- An Interview with Francis Collins." *New England Journal of Medicine* 361(14): 1321-1323.

Usher, A. P., 1929, *A History of Mechanical Inventions*. McGraw-Hill, New York.

Vermeulen, N., Parker, J. N., and B. Penders (2010) "Big, small or mezzo?" *EMBO Reports* 11(6).

Woodward, C. (2011) "National Institutes of Health seek to speed up therapeutic innovations," *Canadian Medical Association Journal* 183(2): E91-E92.

Zaringhalam, M. (2016) "Failure in Science Is Frequent and Inevitable--and We Should Talk More about It," *Scientific American Blog* (https://blogs.scientificamerican.com/guest-blog/failure-in-science-is-frequent-and-inevitable-and-we-should-talk-more-about-it; accessed December 5, 2017).

Table 1: Edge Funding Ratios by Idea Type and Time Period.

(1a) Idea Type	(1b) Characteristic of Idea Type	(2a) Number of Contributions	(2b) Share NIH Funded	(2c) Number of Novel Contributions	(2d) Share NIH Funded	(3) Edge Funding Ratio	(4a) Edge Funding Ratio If Funding Rates Same Across Idea Types	(4b) Edge Funding Ratio If Funding Rates Same Across Research Areas	(4c) Edge Funding Ratio If Funding Rates Same Across Idea Types and Research Areas	(5a) Edge Funding Ratio for 1990-1999	(5b) Edge Funding Ratio for 2000-2009	(5c) Edge Funding Ratio for 2010-2014	(5d) Edge Funding Ratio for 2015-2016
All Idea Types		8637781	50%	653054	59%	1.26	1.05	1.14	1.04	1.12	1.26	1.26	1.22
Gene or Genome	Basic Science	320281	67%	109222	77%	1.40		1.31	1.31	1.12	1.36	1.40	1.41
Amino Acid, Peptide, or Protein	Basic Science	376784	67%	100979	76%	1.36		1.30	1.30	1.16	1.36	1.36	1.34
Cell Function	Basic Science	180244	70%	7610	82%	1.30		1.21	1.21	1.07	1.20	1.28	1.35
Research Activity	Basic Science	261108	43%	6465	51%	1.19		1.16	1.16	1.05	1.18	1.16	1.30
Genetic Function	Basic Science	146435	73%	3476	78%	1.16		1.15	1.15	1.08	1.13	1.15	1.17
Cell Component	Basic Science	93224	73%	8445	82%	1.15		1.11	1.11	1.03	1.11	1.15	1.20
Nucleic Acid, Nucleoside, or Nucleotide	Basic Science	95267	67%	15517	75%	1.15		1.12	1.12	0.99	1.07	1.15	1.17
Molecular Function	Basic Science	205683	69%	18942	72%	1.15		1.10	1.10	1.02	1.14	1.15	1.07
Cell	Basic Science	118015	73%	8744	77%	1.07		1.04	1.04	1.03	1.02	1.07	1.08
Molecular Biology Research Technique	Basic Science	76648	62%	28774	66%	1.02		1.00	1.00	0.92	0.88	1.00	1.11
Organic Chemical	Basic Science	100311	48%	18879	48%	0.98		0.96	0.96	0.95	0.98	0.97	1.00
Laboratory Procedure	Clinical	295296	56%	23798	63%	1.19		1.15	1.15	1.09	1.17	1.19	1.04
Disease or Syndrome	Clinical	283706	44%	19948	46%	1.07		1.00	1.00	1.12	1.18	1.09	0.91
Diagnostic Procedure	Clinical	176915	42%	8563	42%	1.04		1.07	1.07	0.91	1.25	1.03	1.02
Neoplastic Process	Clinical	101140	47%	8728	46%	0.96		0.95	0.95	0.95	0.85	0.95	1.07
Finding	Clinical	494941	42%	12302	40%	0.96		0.99	0.99	0.96	1.05	0.93	1.05
Body Part, Organ, or Organ Component	Clinical	86358	54%	1760	48%	0.91		0.84	0.84	1.00	0.96	0.90	1.00
Clinical Attribute	Clinical	221538	44%	7374	38%	0.91		0.93	0.93	0.94	0.96	0.91	0.81
Pharmacologic Substance	Clinical	349900	51%	69995	44%	0.88		0.87	0.87	1.07	1.03	0.88	0.84
Health Care Activity	Clinical	161458	31%	5754	27%	0.88		0.97	0.97	0.65	0.89	0.84	1.03
Therapeutic or Preventive Procedure	Clinical	313637	36%	17673	29%	0.77		0.81	0.81	0.73	0.77	0.76	0.74
Quantitative Concept	Miscellaneous	301339	43%	5406	60%	1.46		1.33	1.33	0.84	1.18	1.46	1.53
Functional Concept	Miscellaneous	241393	50%	1680	69%	1.41		1.26	1.26	0.96	1.12	1.41	1.20
Biomedical Occupation or Discipline	Miscellaneous	104452	47%	14298	64%	1.38		1.21	1.21	0.47	1.23	1.37	1.38
Intellectual Product	Miscellaneous	371079	40%	17833	42%	1.02		1.03	1.03	0.83	0.94	1.02	1.06
Manufactured Object	Miscellaneous	152426	34%	7828	31%	0.95		0.98	0.98	0.96	1.27	0.91	1.20
Qualitative Concept	Miscellaneous	176011	41%	974	35%	0.85		1.01	1.01	0.79	0.76	0.85	0.89
Idea or Concept	Miscellaneous	141492	39%	1121	18%	0.45		0.53	0.53	0.53	0.60	0.43	0.50
All Other Idea Types		2515573	47%	100122	42%	0.89	0.90	0.89	0.93	0.97	1.03	0.89	0.80

Notes to Table 1: Numbers in Columns 2a-2d, 3 and 4-4c are calculated based on biomedical research articles published during 2010-2016. Please see the main text for explanations and interpretations of each column.

Table 2: Edge Funding Ratios by Research Area.

Idea Type	Characteristic of Idea Type	All Journal Categories	Allergy and Immuno-logy	Behavio-ral Sciences	Bioche-mistry	Biology	Cardio-logy	Cell Biology	Chemist-ry	Environ-mental Health	General Surgery	Medicine	Micro-biology	Molecu-lar Biology	Neo-plasms	Neuro-ogy	Pediat-rics	Pharma-cology	Physio-logy	Psychi-atry	Psycho-logy	Public Health	Vascular Diseases	Zoology	All Other Journal Catego-ries
All Idea Types		1.26	1.15	1.10	1.51	1.22	1.07	1.06	1.19	1.17	1.08	1.28	0.95	1.12	1.13	1.10	1.10	1.10	1.20	1.04	1.03	0.97	1.18	1.18	1.25
Gene or Genome	Basic Science	1.40	1.18	1.23	2.18	1.56	1.54	1.10	1.38	2.69	1.14	1.47	1.60	1.41	1.31	1.16	1.25	1.36	1.30	1.23	1.42		1.22	1.27	1.40
Amino Acid, Peptide, or Protein	Basic Science	1.36	1.26	1.24	1.77	1.30	1.81	1.07	1.23	1.71	1.19	1.48	1.41	1.24	1.22	1.19	1.38	1.22	1.25	1.30			1.52	1.16	1.42
Cell Function	Basic Science	1.30	1.14			1.21		1.10			1.10	1.29	1.27	1.20	1.41	1.11		1.22	1.06				1.29	1.13	1.38
Research Activity	Basic Science	1.19							1.31		1.06	1.15	0.87	1.11	1.28			1.14				1.11		1.18	1.16
Genetic Function	Basic Science	1.16				1.26		1.03			1.09	1.11		1.20	1.07									1.15	1.12
Cell Component	Basic Science	1.15	1.13			1.25		1.07	0.97		1.05	1.10	1.33	1.14	1.02	1.05			1.07					1.10	1.20
Nucleic Acid, Nucleoside, or Nucleotide	Basic Science	1.15	1.05			1.27	0.98	1.00	1.33		1.03	1.08	1.77	1.04	1.08	1.04		1.05	1.02				1.01	1.10	1.17
Molecular Function	Basic Science	1.15	1.15			1.19	1.01	1.02	1.17		1.07	1.11	1.15	1.13	0.97	1.03		0.98	1.04				1.02	1.07	1.17
Cell	Basic Science	1.07	1.12			1.10	1.04	0.95			0.97	1.03		0.91	1.10	0.84							0.98	1.03	1.10
Molecular Biology Research Technique	Basic Science	1.02	0.94			0.98		1.00	0.98		1.03	0.98	0.79	1.01	1.03	1.04		1.01	0.83				0.92	0.99	1.07
Organic Chemical	Basic Science	0.98	1.03			1.21	0.68	1.11	0.90	1.71	1.01	0.94	0.83	1.11	1.01	1.08		1.02	1.56				0.87	1.03	0.91
Laboratory Procedure	Clinical	1.19	1.14		1.50	1.24	0.95	1.06	1.38	1.32	1.11	1.24	0.84	1.07	1.03	0.93	1.35	1.01	1.13	1.16			1.11	1.04	1.28
Disease or Syndrome	Clinical	1.07	0.81		0.83	1.06	0.90	1.05	1.25		1.06	0.99	0.84	1.10	1.10	1.17	0.91	0.94	0.96	1.16			0.82	1.05	1.11
Diagnostic Procedure	Clinical	1.04			1.57		0.91					0.98			0.74	0.95	1.20						0.89	1.20	1.12
Neoplastic Process	Clinical	0.96			1.12			0.69			1.09	1.03		0.92	1.00	0.95	1.34							1.03	0.93
Finding	Clinical	0.96	0.80	0.80	0.95		0.64				1.00	1.06		1.00	1.17	0.88	0.96	0.77		0.99	0.89	1.07	0.85	1.18	0.98
Body Part, Organ, or Organ Component	Clinical	0.91			0.33											0.89									0.94
Clinical Attribute	Clinical	0.91			0.96		0.76				1.04	1.02	0.45		0.95	0.83		0.91						1.06	0.97
Pharmacologic Substance	Clinical	0.88	0.78	0.89	1.50	1.10	0.57	0.99	1.29	1.10	0.93	0.85	0.90	1.02	0.87	0.75	0.84	0.71	1.07	0.54		0.83	0.62	1.07	0.86
Health Care Activity	Clinical	0.88										0.89			1.01		0.75			0.92		0.86			0.91
Therapeutic or Preventive Procedure	Clinical	0.77	0.80		0.85		0.48					0.95		1.01	0.60	0.76	0.94	0.79	0.45	0.89		0.99	0.50	0.98	0.94
Quantitative Concept	Miscellaneous	1.47		1.14							1.16	1.52		1.13	1.28	1.09	1.60			1.16	1.51			1.46	1.47
Functional Concept	Miscellaneous	1.41									0.99		1.41	0.95	1.38									1.16	1.59
Biomedical Occupation or Discipline	Miscellaneous	1.38	1.05			1.14	1.66	1.02	1.21		1.02	1.40	1.03	1.01	1.24	1.35		1.53	1.16				1.41	1.11	1.51
Intellectual Product	Miscellaneous	1.02	0.79	1.18	1.19		0.91		1.66		1.08	1.06		0.96	0.96	0.77	1.05	1.04		0.83	0.79	0.90	0.77	0.99	1.07
Manufactured Object	Miscellaneous	0.95							0.51	0.92	0.66	1.08		0.88			1.30	1.48				1.31		0.71	1.04
Qualitative Concept	Miscellaneous	0.85																							0.87
Idea or Concept	Miscellaneous	0.45																							0.46
All Other Idea Types		0.89	1.06	0.82	1.11	0.63	0.58	0.98	1.10	0.76	0.95	0.87	0.63	0.83	1.06	0.96	0.96	1.13	0.91	0.84	0.84	0.78	0.65	0.89	0.91

Notes to Table 2: Numbers in all columns are calculated based on biomedical research articles published during 2010-2016. Please see the main text for explanations and interpretations of each column.