Does more money lead to more innovation? Evidence from the life sciences.*

Marc Lerchenmueller†
Yale University

Working Paper, July 2018

Abstract: Large sums are often invested into scientific innovation – the creation of new knowledge through scientific research. In this paper, we argue that increasing investments may lead scientists to pursuing average (but more certain) as opposed to higher risk projects, with negative consequences for scientific innovation. Exploiting an exogenous multi-billion dollar shift in the budget of the world's largest financier of scientific research, the U.S. National Institutes of Health (NIH), we find that the influx of more money led to a significant decrease in scientists' innovation productivity (14% fewer papers published), its significance (16% less citations generated), and novelty (9% reduction in unprecedented content). These negative effects become more pronounced when we, in addition to the macro level (federal budget), also account for differences in funding at the micro level (project budget). The decrease in scientific innovation is primarily driven by top scientists changing their research strategy with greater funding availability; the data reveal a 1.7x to 3x larger reduction in scientific innovation for scientists in the top quintile versus scientists in lower quintiles of the capability distribution. We conclude with implications for public policy, corporate strategy, and the allocation of resources in support of scientific innovation.

*I thank the Initiative on Leadership and Organization at the Yale School of Management for generous financial support. I also received financial support through a grant from the German Research Foundation (DFG). I thank seminar participants at the Health Care Policy Department of Harvard University, Mannheim Business School, and the Yale School of Management for valuable feedback, as well as participants at the workshop on the Organization, Economics, and Policy of Scientific Research at Bath and at the Druid conference at Copenhagen. I am also grateful for comments and suggestions by David Audretsch, Karin Hoisl, Ingrid Nembhard, and Olav Sorenson on earlier versions of this work. The usual disclaimer applies.

†Yale School of Management, 165 Whitney Avenue, New Haven, CT, 06511; marc.lerchenmueller@yale.edu

1 Introduction

Few would question the idea that scientific innovation[1] fuels economic growth and improves human well being (DeLong, 2000; Mokyr, 2010). It may therefore also seem self-evident that corporations, and societies more broadly, are willing to commit a lot of money to scientific innovation. The latest estimates suggest that total U.S. R&D investments have reached an all-time high of $499 billion, more than a two-third increase in current dollars since the early 2000s (Boroush, 2016). But there is more to such expanding investments than meets the eye, as the way in which the money is spent also changes profoundly across OECD countries. Corporations increasingly focus on downstream commercialization of research (the "D" in R&D) across most industries, while delegating the science (the "R" in R&D) that is often picked up by research universities that are, in turn, backed by corporate and even more so public funding (Arora et al., 2015). In the life sciences, public funding through the U.S. National Institutes of Health (NIH) has risen to similar levels as corporate R&D funding, with NIH money flowing almost exclusively to basic research whereas only about one out of five corporate R&D dollars still does (Moses et al., 2015). The deepening rift in investments is not explained by corporations shedding a fruitless activity; to the contrary, scientific innovation continues to drive competitive advantage as indicated by, for example, corporate patents' high citation rates of academic publications across industries (Arora et al., 2015; Li et al., 2017). On the face of it, the increasing segregation of "R" and "D" may then reflect effective division of labor if universities are better at innovative research and corporations are better at its development.

In this paper, we investigate whether increasing the amount of money that flows to university research increases scientific innovation? Providing evidence towards answering this question is important because, unless public funding of academic research can make up for the decline in corporate research, we may experience reduced economic growth and corporate and public funders alike may have reason to revisit their innovation strategies[2]. To forward our theoretical understanding of how money shapes the conduct of scientific research, we turn to the broader literature

[1] We refer to scientific innovation as the creation of new knowledge through mostly basic research, and we use the terms science, scientific research, and basic research interchangeably, mindful of the empirical difficulty in separating these concepts.

[2] Prior research has mostly focused on how different public and private funding mechanisms identify promising research proposals through peer review (Azoulay et al., 2011; Boudreau et al., 2016; Li and Agha, 2015) and how subsequent grant money receipt influences individuals' productivity (Benavente et al., 2012; Jacob and Lefgren, 2011; Whalley and Hicks, 2014). By contrast, our study seeks to determine how the influx of more money may *change* funding decisions, scientists' research strategies, and ultimately scientific innovation.

on the economics of innovation and examine the relative likelihood of three possible dynamics. We start with thinking about innovation as a knowledge creation process that uses a set of inputs to generate outcomes. Scientific innovation is generally performed by well paid scientists and ever larger teams of scientists (Wuchty et al., 2007). The infrastructure needed to conduct scientific research has also become more expensive (Stephan, 2012). From this perspective, an increase in available funding might have a positive effect on research outcomes. Yet changes in the amount of money available for investment might also change the knowledge creation process itself. On the one hand, financiers may shift knowledge creation to more high-risk projects because having more money for investing allows for better risk diversification and makes financing of higher risk projects more feasible and attractive. Since high-risk projects tend to yield research outcomes of great value, a few success cases may already compensate for many other failures (Fleming, 2007). This scenario would again suggest a positive relationship between money and scientific innovation. On the other hand, scientists may change their research strategies – one way or another – as more money becomes available. We find that standard economic analysis is largely silent on how money may shape the motives of scientists performing the research. We therefore also look to literature on how competition for resources shapes motives for knowledge creation (March, 1991).

To test the relative applicability of the outlined dynamics, we exploit an exogenous multi-billion dollar shift in the budget of the world's largest financier of scientific research, the NIH, at the turn of the century. This setting provides an unusual opportunity for observing both sides of the innovation table – funders and innovating scientists. We use the NIH budget doubling and a matched control group design to address our empirical task of disentangling the marginal impact more funding might have on scientific innovation from selection into the funding scheme. Our data consists of two sets of matched scientists. The first set includes scientists who received their first major research project grant – the R01 grant that accounts for more than half of all NIH grant dollars – prior to the NIH budget doubling, matched to observationally equivalent[3] scientists who did not yet receive their first R01 grant. The second set contains scientists who received their first R01 grant towards the end of the budget doubling and respectively matched scientists. These sets combine to 14,547 scientists[4]. The data allow us to estimate the effect of the

[3]We match scientists based on their relative performance across the five R01 evaluation criteria (see *4. Data*).

[4]We use Coarsened Exact Matching (CEM), matching two control scientists to one R01 recipient, on average. Our results remain significant for other matching ratios and matching techniques. See also section *5. Results.*

first R01 grant on scientists' subsequent innovative output in terms of productivity (subsequent papers published), its significance (citations generated), and novelty (unprecedented combinations of scientific content). Moreover, we can estimate whether the "R01 effect" differs conditional upon the plausibly exogenous assignment of would-be first time R01 scientists to NIH budget states, that is, as more money becomes available to fund more projects and to fund projects with larger sums as a result of the NIH budget doubling.

We find that the greater availability of funding with the NIH budget expansion is associated with 14% fewer publications within five years post grant award, with 16% fewer citations on these publications, and with a 9% reduction in the research's novelty among scientists who received their first R01 grant. This relative decrease in scientific innovation becomes more pronounced when we, in addition to funding differences at the macro level (increase in the NIH budget), also account for funding differences at the micro level (project budget). Note, the annual R01 grant size increased by about $40,000 (or 16%), on average, in the wake of the NIH budget doubling. Accounting for project-level funding, the previously identified effects increased to 20% fewer publications, 21% fewer citations, and 10% less novelty. The data further discount the possibility that these negative effects are driven by the financier through allocating the additional money to projects of systematically lower quality. Instead, scientists' research strategies change with more money being available. In particular, top scientists engage in less novel but more certain (average) research, thereby driving the decrease in scientific innovation as funding expands.

In addition to the theoretical implications of these results, our study holds value for management practice in corporate and public settings. As channeling more money for research to universities might not lead to commensurate gains but even losses in scientific innovation, corporations may need to bring back some basic research. This likely would also require withstanding short-termism and again taking a longer term view on innovation and corporate growth (Barton et al., 2017). For both managing corporate and public research, our findings highlight the importance of engineering competition for funding so that performing just better than average becomes less attractive relative to pursuing higher risk-reward research in service of innovation. Lastly, the current discourse on contracting NIH funding notwithstanding, the reallocation of R&D money across OECD countries, and the recent formidable capital injections into university research in Canada, China, and Japan point to the practical relevance of our findings on a global scale.

2 Innovation as a process of knowledge creation

Innovation is the economic activity that can most closely be conceptualized as the creation of new knowledge (Audretsch and Feldman, 1996). Knowledge creation requires a set of inputs, chiefly among them are physical capital, like laboratory space (infrastructure) and centrifuges (technology), and human capital (Becker, 1994; Nonaka, 1994). More than half of corporate R&D spending is the wages and salaries of highly educated scientists and engineers (Hall and Lerner, 2010). In the academic sciences, investment into innovation is even more tied to people as a virtue of financing mostly research; the "D" in R&D plays a subordinate role in academic settings.

Across public and private sectors, increasing research and development expenditures can be traced back to profound changes in the way scientific research is conducted. Since the 1980s, we observe an evolving dominance of teams of scientists creating new knowledge rather than individuals. This phenomenon is especially pronounced in resource intensive disciplines (physics, engineering, life sciences) where the evolution towards teams appears exponential relative to, for example, the arts and humanities where the development is fairly flat (Wuchty et al., 2007). Part of the rational for creating knowledge in teams may therefore be the ability to share increasing costs for the physical assets needed for research. Consider as an extreme example the creation of the $8 billion Large Hadron Collider at the CERN in Switzerland which required teams of several thousand physicists and engineers to create and countries to split the bill (Bikard et al., 2015). It is neither uncommon that firms justify their engagement in research joint ventures from sharing expensive technology.

Besides higher prices for physical assets, it seems that greater intellectual demands in the quest for new knowledge drive up the size of teams and the cost of research. A longstanding stream of literature conceptualizes knowledge creation as a recombinant search across idea spaces (Gilfillan, 1935; Schumpeter, 1939; Basalla, 1988; Weitzman, 1998). New knowledge results from the combination of new with existing ideas or from the recombination of existing ideas in novel ways (Kogut and Zander, 1992). Instead of a trial-and-error search for combinations, scientific innovation is characterized by using theories that inform which combinations may be most promising given a certain goal (Fleming and Sorenson, 2004). In that sense, scientists may be viewed as cartographers of idea spaces that can use their mental maps as guidance to the most promising solutions. The

problem now is that these mental maps become very large. *PubMed*, the core bibliographic reference in the life sciences, adds more than 300,000 scientific articles to its repository on a yearly basis, for example. Burgeoning knowledge triggers increased specialization among scientists (Jones, 2009). Specialization helps in navigating the idea space locally, but that oftentimes precludes making more distant combinations that tend to be associated with greater innovative value (Hargadon and Sutton, 1997; Fleming, 2001). In turn, limitations from specialization seem to have contributed to the call for more interdisciplinary research that increasingly spans specializations, disciplines, and even countries (Jones et al., 2008). In a nutshell, research teams appear to grow larger in size to access needed specialized knowledge while also drawing on diverse knowledge to improve the odds of scientific innovation.

Against this backdrop, having more money to fund research would appear to promote scientific innovation. Trivially, more money allows for assembling larger and better teams that can exploit benefits from specialization and diversity. Funding better specialists with expertise in a specific area should allow for faster and more accurate identification of promising ideas in a local space (Fleming and Sorenson, 2004). And if local spaces become too large and too complex to navigate, even for groups of scientists, money allows for the procurement of supporting resources. For example, consider high-throughput screening (HTS) in the fields of biology and chemistry that use robots, data processing algorithms, and sensitivity detectors to conduct millions of chemical or genetic tests that allow for the identification of active compounds or genes that modulate a particular biomolecular pathway that scientists consider of interest. As to the benefits of financing team diversity, we recite the long established argument that collaboration across different areas of expertise should increase the potential combinatorial opportunity for creating new knowledge (Gilfillan, 1935; Weitzman, 1998). To the degree that more money brings together better technology and better teams of scientists, we would expect the average research project to yield better outcomes.

What is left open is whether the expected improvement in the average outcome goes in hand with an increase or decrease in the variability of research outcomes. Several arguments and empirical evidence suggest that devoting more resources tends to improve the reliability of research outcomes. Technology becomes more efficient in helping with minimizing uncertainty in the research process. High-throughput screening does not only help with identifying promising avenues for research, it also eliminates a plethora of alternative paths that would have otherwise clouded the way. In

doing so, the likelihood of low value research (failure at the extreme) should be reduced. But it seems at least questionable whether investments into infrastructure or technology materially affect the likelihood of outstanding scientific innovation. As to the effect of teams on variability, financing more knowledgeable people should also make for a more rigorous selection process of ideas, thereby reducing the likelihood of low innovation value or failures. Diverse backgrounds reduce groupthink (Janis, 1972) and allow for evaluating ideas from different perspectives, in turn, increasing the likelihood that only better ideas stand a chance for further pursuit. Analyzing over half a million patented inventions, Singh and Fleming (2010) documented that teams of innovators were significantly more likely to trim poor outcomes, while the effect on creating innovation of particularly high value was less pronounced.

Taken together, the preceding arguments and related empirical evidence would point in the direction of greater investment leading to improved research outcomes, on average, as well as to a tangibly lower likelihood of failure. Both of these effects are illustrated with stylized probability density functions of scientific innovation, and a shift from the blue to the red distribution with more money financing better resources (Figure 1a). To the degree that investing more money leads to greater reliability in research outcomes from a reduction in the left-hand tail, it remains a possibility that an increase in the average research outcome coincides with a decrease in the right-hand tail – a lower likelihood of scientific innovation of high value. The dashed vertical lines mark the top 20% of scientific innovations to visualize this scenario. Innovation of high value tends to arise from the new combination of ideas and that process also entails higher risks (Fleming, 2001). If more money may either lead financiers to pick more high-risk projects, or lead scientists to alter their research strategies, we would expect the red distribution to look differently than stylized in Figure 1(a). Up to here we have implicitly assumed, however, that the opportunities for knowledge creation (projects proposed or financed) remain constant and we now turn to how these opportunities might change with more money becoming available.

2.1 More money, funding strategies, and knowledge creation

Innovation projects have characteristics that differentiate their funding from other investment decisions. For starters, it often requires a long time (sometimes several years) to execute an innovation project, which poses challenges to financing (Hall and Lerner, 2010). Moreover, the potential of

innovation projects tends to be difficult to assess prior to actually embarking on the lengthy journey. The inherent uncertainty has implications for how to decide which projects to finance (Nanda and Rhodes-Kropf, 2017). These features are independent of whether the financier and innovator are the same person or entity, but in practice they are most often not, adding yet another layer of complexity. A corporation that conducts R&D is answerable to financing shareholders and so is the innovating scientist to his or her manager or scientific funding body. In light of these characteristics, we discuss two arguments that would suggest a positive effect on scientific innovation from more money being available for funding.

To begin, money and innovation are related through the fact that any financing decision of an innovative project involves an evaluation of the risks against potential outcomes. Aiming for higher degrees of innovation entails higher risks (including failure and sunk investment), whereas more incremental innovation can be had with more certainty (and outcomes of lower value in expectation) (Singh and Fleming, 2010; Dewar and Dutton, 1986). Past research suggests several ways in which greater availability of money influences this trade-off to the benefit of innovation. Starting with a cross-sectional view, more money should enable the financing of more projects across which risk can be diversified. In theory, more high-risk projects can be funded because more money is available to simultaneously finance lower risk projects that compensate for expected losses from the former. Yet since high-risk projects also tend to generate outsized value, a few successes are often enough to outweigh losses, likely leading to a net positive effect on innovation (Fleming, 2007). According empirical evidence spans fields and units of analysis, from economic geography and regional innovation clusters to strategy and firm-level R&D portfolios. For example, Klingebiel and Rammer (2014) found that investing into a broader range of innovation projects had positive performance effects for firms intending to create relatively more novel products by departing further from their knowledge base (i.e., that accept higher risk).

Recent research that takes a dynamic and longitudinal perspective has begun to also suggest a positive relationship between more money being available for financing and innovation. Consider that most innovative projects are funded through a series of investments as opposed to a lump sum at the start. The life sciences are a case in point as more than half of NIH research project grants seek a renewal at least once, summing to at least eight to ten years of investment (Lauer, 2014). As innovative projects progress, uncertainty tends to subside, making it easier for financiers to evaluate

whether additional investments are warranted with the passing of time and more information becoming available. Said differently, the optimal investing strategy often has an option-like character (Hall and Lerner, 2010). The option template for investing, however, tends to not fit with highly innovative projects which carry so much uncertainty at the outset that a lot of time and sometimes a certain level of trial-and-error is needed to progress. If funding is constrained, investors may even shy away from initial financing in anticipation that the available money is insufficient to fund the project through additional stages, in turn, making it more likely that any initial investment is lost (Nanda and Rhodes-Kropf, 2017). As such, the funding of projects with high innovation potential may especially hinge on more money being available for investment. Consistent with this logic, Howell (2017) found that the U.S. Department of Energy grant program spurs innovation in cleantech by financing relatively inexpensive yet otherwise difficult to finance proof-of-concept studies that, if successful, attract more money from other investors.

Figure 1(b) visualizes these dynamics. To the degree that more money leads investors to financing more research of higher risk, the likelihood of scientific innovation of high value increases as indicated by the shift from the dashed blue to the dashed red line (the 80^{th} percentiles of the respective distributions). This positive effect, however, comes at the cost of a higher risk of failure which should, at least in part, be compensated by funding research that diversifies risk (the left-hand tail still thickens but less so than the right-hand tail). Combined, we would expect the average research outcome to improve.

2.2 More money, research strategies, and knowledge creation

A feature that is largely absent from the standard economic analysis of innovation (including from the preceding arguments) is that the research leading to scientific innovation is performed by individuals that have their own set of motives and preferences (e.g., Gilbert, 2007). Agency theory is the branch of economics that is concerned with the analysis of agents (here innovators) acting on behalf of a principal (financier), and the latter has incomplete control over the actions of the former (Eisenhardt, 1989). A central pillar of this vast literature is about how principals can design incentives in a way to motivate the agent to act in the best interest of the principal. The focus here is again on the financier, but this perspective incorporates potential motives of the innovator.

Seminal work on principal-agent conflicts dates back to the 1970s (Eisenhardt, 1989), yet re-

search that looks at incentive design in the context of innovation increasingly appeared over the past few years. These studies suggest that optimal incentive schemes for innovation exhibit a high tolerance for failure and reward for long term success (Manso, 2011; Ederer and Manso, 2013). One would expect that implementation of such incentive schemes requires deep pockets on part of the financier. Similar to the arguments extracted from literature on how funders go about picking projects, if more money is available for funding then financiers can stick longer with a project (tolerating failure in earlier stages) and can take on higher risk projects with higher scientific innovation potential in the long run. The addition from agency theory would be that innovators can also be motivated to propose and execute projects accordingly. Empirical evidence, though, is rather thin. One study that contrasted NIH funded scientists to scientists funded by the Howard Hughes Medical Institute (HHMI) that explicitly commits funding long term, provided evidence that HHMI investigators do take on riskier projects resulting in greater scientific innovation (Azoulay et al., 2011). Note, that HHMI investigators also receive between two and three times more money than NIH R01 investigators. Importantly, though, HHMI intends to fund people and not projects, which alters the dynamics of competition for funding[5].

Competition for funding pits scientists against each other, whether in academia or in corporate R&D internal markets. The NIH fund allocation process, with its commitment to financing projects according to the relative position of an application on a percentile ranking from 1 to 99[6], exemplifies competition for relative rank. March (1991) highlighted that relative position is increasingly affected by variance (the risk being taken) as competition increases. Meanwhile, risk taking has a negative effect when one seeks to avoid low positions in the ranking. Given a certain level of competition, scientists can pursue research that may make them stand out or doing just better than average.

This line of reasoning still assumes that scientific innovation results from several individual research outcomes that are independent draws from a distribution and that scientists do not have much influence over the distribution itself. But as one scientist's research strategy should depend on the strategies of competing scientists, it seems more reasonable that scientists actively decide on their research program. Going back to the notion of scientific innovation being a recombinant

[5]HHMI funds less than 100 investigators, while even first time R01 awards are allocated to about 2,000 scientists per year.

[6]An application's percentile score represents the percentage of applications from the same study section (area of research) and reviewed in the same year that received a better evaluation.

search, it stands to reason that scientists who achieve better research outcomes tend to have a better grasp of the space across which ideas can be combined. Put differently, they should have a better sense for what type of research is needed to meet expectations for funding when compared to scientists of lower capability. While the latter group of scientists may be more challenged in influencing their expected research outcome in a reliable fashion, they can decide on the risk they are willing to take. When competition for funding is fierce, one would expect scientists at the lower end of the distribution to take higher risks to stand a chance against better scientists as only higher relative ranks get funded (March, 1991). The punishment for failure approaches zero as the likelihood of reaching the top ranks was low to begin with, while taking high risks may entail a high return (grant funding). This behavior, in turn, forces better scientists to also take risks.

As competition for funding lessens with more funding being available, however, the incentive for higher risk research diminishes from both ends of the spectrum. It would seem rational, though, that scientists who are better able to adjust their expected outcome will adjust their research strategy the most. As aforementioned, funding rates increased by one third in the wake of the NIH budget doubling to a level where about one in three investigator proposals were funded (Rockey, 2014). Scientists toward the top end of the distribution would have less to gain from pursuing higher risk as opposed to more incremental research that still leads to funding when compared to a situation where only say the top ten percent of projects are funded. On the positive side, such risk adjustment should also reduce the likelihood of failed research (or research of extremely low value). But considering the rich literature that documents the outsized contribution of top scientists to research outcomes, from corporate R&D programs (Rothaermel and Hess, 2007), to public-private research partnerships (Zucker and Darby, 1996), to the academic sciences (Merton, 1973), we would expect a greater decrease in the likelihood of high value scientific innovation that is not compensated by avoided failures (a stronger thinning of the right versus left-hand tail). Figure 1(c) depicts these dynamics and the ensuing deterioration in average research outcomes.

3 Context and empirical strategy

Building from our review of the literature, we have argued that innovators' responses to changes in available funding have received relatively little attention. Generally speaking, empirical research

on the economics of innovation has focused on the analysis of R&D spending from a financier point of view, at least partly due to data availability and measurement challenges (Hall and Lerner, 2010). Our empirical focus on the life sciences uniquely provides us with a view into funding decisions by financiers *and* scientists deciding what research to pursue. NIH funding is vital to U.S. biomedical research with over 80% of basic life science laboratories receiving some form of NIH funding (Sampat and Lichtenberg, 2011). Although the NIH administrates many different grant programs, the largest and most established is the R01 research project grant, which constitutes more than half of all NIH grant funding at any given time and serves as the primary funding source for U.S. principal investigators and their laboratories (Li and Agha, 2015). Since different grant mechanisms may have different funding criteria, and since the R01 grant clearly stands out, we focus our attention on this funding mechanism.

A centerpiece of our research design is using the NIH budget doubling as a natural experiment to estimate the causal effect more money might have on scientific innovation. For context, the U.S. Senate endorsed the goal of doubling NIH's budget within five years in 1997. Starting already in 1998, the U.S. Congress increased NIH appropriations from $23 billion to $39 billion by 2003[7] (Johnson, 2016). Note, NIH investments into scientific research came then close to pharmaceutical firms' investments into R&D, although only about one out of five dollars flows to basic science in case of the latter (Moses et al., 2015). While a few studies have looked into how the NIH budget doubling affected the composition of the biomedical research workforce (e.g., Blume-Kohout and Clack, 2013), surprisingly little attention has been devoted to the effects on research. The one quantitative assessment we are aware of found an increase in the total number of publications, but that increase appeared not as different as one might expect when compared to research output in other countries that did not experience such an influx of grant money (Sachs, 2007). Importantly, though, studying the effect of the budget doubling at such an aggregate level may provide directional results at best and mostly noise in the worst case. For example, medical schools committed $15 billion to new research facilities between 1998 and 2007 compared with $3.2 billion prior to the budget doubling (1990 to 1997) (Stephan, 2012). To put this infrastructure to use, applications to the NIH shot up at the end of the budget doubling[8] (Zerhouni, 2006). It was especially established

[7] Constant 2015 dollars

[8] Remarkably, applications rose to an extent that overall success rates actually began to decline and had reached pre-budget expansion levels by 2007.

investigators who accounted for the flurry; the share of researchers with more than one R01 grant grew by over one third during that time (Stephan, 2012). These dynamics concern us to the degree that they introduce noise that is unique to the context[9] when we seek to analyze general effects more money may have on scientific innovation.

Instead of considering R01 recipients in general, we therefore focus our attention on *first-time* R01 recipients. The NIH defines new investigators as those who apply for their first major research project grant (generally the R01 grant) and evaluates their applications separately from established investigators. This is because new investigators are considered as the engine of innovation who bring new ideas and methods to the research enterprise, deserving dedicated support in service of medical advancement (NIH, 2017a). Oftentimes, the first R01 grant also opens the door to a tenure track career as it finances the caliber of independent research expected from faculty (Lerchenmueller and Sorenson, 2018). A positive shift in the NIH budget would therefore appear to be of particular relevance for scientific innovation in this cohort.

As a preliminary exercise, we looked at how the NIH budget expansion translated into changes in first-time R01 awards and recipients' subsequent research outcomes in aggregate. Figure 2 shows that the number of first-time R01 grants awarded between 1995 and 2004 (green dashed line) increased faster than the overall NIH budget (red dashed line) and began to stabilize midway through the budget expansion. The number of first-time R01 grants increased from about $1,200$ to over $2,000$ per year. The budget per grant also increased from about $260,000 to over $300,000 per year, summing to more than $1 million over the course of a typical four to five years grant cycle (not shown in Figure 2). Both, the total number of publications within five years of first R01 award (one grant cycle), as well as the total citations on these publications increased strongly at the start of the NIH budget doubling. After 2000, the improvements in these aggregate measures flattened out, suggesting potential diminishing marginal returns from additional funding increases. To assess the novelty of these publications, we counted instances of never before combined article keywords (called MeSH terms in the life sciences; see also section 4 for details). Both the number of publications with at least one novel keyword combination (green solid line) as well as the total number of unprecedented keyword combinations (solid yellow line) declined after an initial expansionary phase.

[9]For example, senior scientists scaling up research programs for reasons besides scientific innovation may or may not be a concern when more money is invested in research and development in other contexts.

3.1 Empirical strategy

To evaluate the causal effect that more money might have on scientific innovation we must, however, address a fundamental inference problem. For a given scientist within a given institutional environment, one cannot observe the counterfactual impact his or her research would have, had it been funded with more as opposed to less money. Ideally, one would randomly assign different levels of funding to individuals of identical (or highly similar) capabilities and examine subsequent research output. While one cannot replicate this ideal experimental design, we develop an empirical strategy that takes advantage of the exogenous shift in the NIH budget to isolate the marginal impact of additional funding on scientific output from the effect of selection into the funding scheme.

Our approach exploits two key elements of our setting. First, all publications that result from research supported by NIH grants have to acknowledge the supporting grant by federal law. We can therefore assess the impact of greater funding by comparing first-time R01 grants awarded after to those awarded prior to the budget expansion and examine patterns in the resulting number of publications, citations, and the research's novelty[10]. Second, this setting allowed us to construct a data set that not only includes the post-grant publication histories of first-time R01 recipients but also their pre-grant publication histories that we can use to identify observationally equivalent scientists with respect to R01 grant evaluation criteria but who did not yet receive R01 funding. Specifically, we assemble two sets of scientists. The first set includes first-time R01 recipients during a three-year window (1995 to 1997) prior to the NIH budget doubling decision, matched to non-R01 scientists with similar publication records up to the end of our observation window. The second set identifies first-time R01 recipients again during a three-year window (2001 to 2003)[11], but three years after the budget doubling decision and respectively matched non-R01 scientists (see section 4 for details).

The focus on first-time R01 scientists is important for our empirical strategy in yet two additional respects. Within the group of first-time R01 applicants (i.e, new investigators), the NIH encourages applications within ten years of the terminal degree (MD/ PhD) by recognizing these individuals as early career investigators (NIH, 2017a). Limiting the selection of both first-time

[10] We formally test and provide evidence for the comparability of first-time R01 cohorts across NIH budget states in the next section.

[11] Our results were similar when we altered observation windows on the margin or narrowed them.

R01 scientists and non-R01 scientists for matching to those individuals whose publication histories do not extend beyond 10 years prior to the respective observation windows (1995-97 and 2001-03) should yield a fairly homogeneous sample. This approach reduces, for example, the likelihood of including scientists who have received substantive funding from other institutions we do not observe[12] prior to their first NIH R01 grant. We additionally limit the identification of non-R01 scientists to those that have received some form of NIH funding other than an R01 grant. This condition further increases the likelihood that we are considering scientists of similar capabilities.

Perhaps even more important, the focus on early career scientists renders their assignment to NIH budget states plausibly exogenous with respect to first-time R01 receipt. Being eligible for early career designation as R01 applicant requires both, not being out of graduate school for much longer than ten years[13] and a publication history worthy of principal investigator status on a million dollar R01 grant. Scientists usually come close to meeting both requirements when they are close to the ten-year mark post graduate school. In our data, scientists had about 8.5 years of publishing experience, on average, prior to applying for their first R01. In other words, it would appear highly infeasible for these scientists to strategically time their first R01 application so that it may fall either prior to or after the NIH budget doubling – recipients in the first window (1995-97) would generally be too experienced four to six years later, while those in the second window (2001-03) would likely have been too inexperienced to be competitive.

Taken together, because we observe research output of scientists who have been awarded with their first R01 grant prior or post the NIH budget doubling decision, and because we are able to identify a counterfactual estimate of scientific output that would have occurred if R01 funding had *not* occurred, we can identify the causal impact of greater funding on subsequent scientific output. We develop an estimation that includes "scientist pair" effects that identify each matched group of R01 and non-R01 scientists, a dummy variable for whether the scientist was observed prior to or after the NIH budget doubling decision (capturing counterfactual scientific output), a dummy variable for all first-time R01 scientists (identifying the marginal impact of funding on scientific innovation prior to the budget doubling), and an interaction effect of the two dummies

[12] Cardiologists, for example, might receive fellow to faculty grants from the American Heart Association (AHA) that serve a similar purpose as the first R01 grant. Yet it would be unusual for scientists to secure substantive AHA and NIH R01 funding within ten years of graduating from a PhD program or medical school.

[13] Certain events, like child birth, make an extension of the ten year mark possible according to NIH policies.

that captures the difference in the impact of funding as funding expanded (our treatment effect of interest). Our estimation strategy therefore follows a difference-in-differences logic whereby we determine the change in the degree to which funding produces scientific innovation.

4 Data

To implement the articulated empirical strategy, we addressed four challenges: (i) We must demonstrate that the requirements for first-time R01 receipt were comparable prior to and after the NIH budget doubling decision; (ii) we needed to construct samples of control scientists with their publication histories to assess counterfactual scientific innovation in the absence of R01 funding; (iii) we must link first-time R01 grants to all publications associated with grant funding and assess different dimensions of scientific innovation for both R01 and non-R01 scientists; (iv) we needed to create samples of grant-publication relationships that can be used to identify the marginal impact of more money being available through the NIH budget doubling.

We assembled data from three core sources to address the outlined challenges. The *NIH ExPORTER* records all NIH funded research projects from 1985 to today (NIH, 2017b). The *PubMed* database is the most comprehensive listing of articles in the life sciences, including more than 25 million articles and more than 70 million authorships from the 1800s to today (National Library of Medicine, 2017). Finally, we used the *Author-ity* database that assigns author IDs to all authors on *PubMed*-listed articles that were published up to 2009 (Torvik and Smalheiser, 2009; Lerchenmueller and Sorenson, 2016). Appendix I provides further details on the author disambiguation that assigns authors across *PubMed* recorded articles with greater than 99% accuracy.

4.1 R01 grant evaluation criteria and the NIH budget doubling

The NIH regulations stipulate five criteria by which R01 applications are to be evaluated. Table 1 provides names and definitions of all variables included in our study, starting with our variables that capture R01 criteria. Since we also use these variables to later match R01 to non-R01 scientists, we label this set of variables matching variables[14].

[14]We used logged specifications of our matching and dependent variables to account for skewed distributions and to enable matching that is consistent with funding effect estimates that can be interpreted as percentage changes in scientific output.

Table 2 provides a descriptive overview of the R01 evaluation criteria and our matching variables. The first criterion, *significance*, essentially asks whether the proposed research by the applicant will progress the field of research. Citations to the applicant's past work serve as an immediate indicator of whether the field tended to make use of the applicant's research. Some have considered citation counts as an article-level metric of research quality but even if other factors influence these counts they most certainly reflect the attention received by the applicant's research and that factor influences grant committees' decisions.

The *investigator* is assessed as to whether he or she had an ongoing record of accomplishments that promises conversion of R01 funding into scientific output, and we used the logged count of prior articles per years of scientist's experience for this criterion.

Third, the potential *innovation* from the proposed research is another explicit evaluation criterion. Extant literature conceptualizes innovation as the combination of new ideas with existing ideas or the recombination of existing ideas in novel ways (e.g., Kogut and Zander, 1992). Accordingly, we identified unprecedented combinations of article keywords (MeSH terms) assigned to scientists' articles and calculated the proportion of novel combinations relative to combinations that existed in the entire *PubMed* database prior to the article's publication. To elaborate, core ideas and concepts in the life sciences are identified by Medical Subject Headers (MeSH terms) that accompany all publications. The MeSH lexicon is a controlled vocabulary used by the U.S. National Library of Medicine to index articles for *PubMed*. Note, MeSH keywords are assigned *not* by authors but by professional science librarians trained specifically to perform this task, eliminating bias that might otherwise result if scientists could influence this keyword assignment (Boudreau et al., 2016). Since not all novel keyword combinations automatically inhere the same innovative value, however, we weighted the proportion of novel keyword combinations by the impact factor of the journal the article was published in. This approach ascribes greater innovative value to novel keyword combinations that appear in *Science* or *Nature*, for example, as opposed to less influential journals. We averaged the thus obtained measure of novelty across scientists' articles to reflect the novelty of scientists' bodies of work.

The last two criteria, *approach* and *environment*, probe whether an applicant in a given research environment is likely to accomplish the proposed research. To approximate the former, we again used the logged count of prior articles per years of scientist's experience because it provides a

direct measure of project conversion adjusted for time. The scientific *environment*, meanwhile, may reflect several aspects. For example, institutions may differ in the quality of the available infrastructure, the time available to scientists for conducting research (versus administrative or clinical duties), and the experience of peers the applicant may draw upon. We argue that the institution's competitiveness in terms of securing R01 funding captures these various aspects well, and we converted the number of R01 grants an institution received per year, on average, into a percentile rank for matching. Finally, past research also suggested the presence of bias (conscious or unconscious) in evaluating first time R01 applicants of different sex (Lerchenmueller and Sorenson, 2018). We therefore also matched on applicant's sex.

Using the publication histories of all scientists prior to potential R01 grant receipt, we ran logistic regressions of the likelihood of receiving the first R01 grant. Table 3 shows that all variables capturing R01 evaluation criteria were significant predictors of R01 receipt prior to (model 1) and after the NIH budget doubling (model 2). In fact, the obtained effect estimates closely resemble prior research published by the NIH (Eblen et al., 2016). Of note, the 95% confidence intervals for all R01 evaluation criteria overlap, indicating that the composition of R01 and non-R01 scientists remained fairly homogeneous across NIH budget states. Moreover, the pooled regression (model 3), including explicit interaction effects of the R01 evaluation criteria with a dummy variable for the NIH budget state (1 if the scientist belonged to the cohort post the NIH budget doubling decision and 0 otherwise), yielded no significant interactions. Together, these results indicated that the caliber of first-time R01 recipients (and non-R01 scientists) was comparable across budget states as well as that our variables used for evaluating the likelihood of R01 receipt are appropriate for matching R01 to non-R01 scientists.

4.2 Counterfactual scientific innovation

We used coarsened exact matching (CEM) to pair first-time R01 grant recipients with observationally equivalent scientists who would appear similarly eligible to receive first-time R01 funding but who did not yet[15]. The fact that the NIH stipulates five clear criteria for evaluating R01 applications, and the fact that our logistic regressions (Table 3) lent credence to our measurement,

[15]Recent research suggests that CEM has several advantages over other techniques that also match on observable characteristics, such as propensity score matching (see Iacus et al. (2012)). Nonetheless, our results were similar when we used, for example, nearest neighbor matching with propensity scores. See also *Table 8*.

facilitate the matching.

The selection of observationally equivalent controls based on our matching variables still involves a trade-off. On the one hand, coarser matching increases the number of available non-R01 scientists for each first-time R01 recipient, thereby generally improving the precision of effect estimates (smaller standard errors). But, this precision comes at the cost of reducing the value of the matching in adjusting for real differences between R01 and non-R01 scientists. On the other hand, finer-grained matching reduces the number of equivalent matches (potentially down to zero leading to exclusion of an actual R01 recipient), but better accounts for variation in the data. The richness of our data allowed us, though, to exactly match non-R01 scientists to R01 recipients across the six criteria (Table 2). Specifically, we determined the distributions of our matching variables and coarsened our data according to the distributions' quintiles for matching. This approach resulted in the match of two non-R01 scientists to one R01 recipient, on average.[16] The final samples for the first set of analysis included 2,051 R01 recipients and 4,098 non-R01 scientists prior to the budget doubling decision and 2,868 R01 recipients and 5,530 non-R01 scientists post the budget doubling decision. We conditioned our models of scientific innovation on the set of matched non-R01 and R01 scientists, thereby controlling for the characteristics of the first-time R01 recipients and for the variables on which the R01s and non-R01 scientists had been matched.

4.3 Scientific innovation

NIH funding statutes require that scientists list grant support in the acknowledgement section of their published articles, with failure to do so punishable by federal law and by potential disqualification from further NIH funding (Lerchenmueller and Sorenson, 2016). The *NIH ExPORTER* records publications (identified with the unique *pmid*) that emanate from grants (identified with a unique grant number), and the connection of the two IDs have therefore high fidelity. When comparing the number of articles of first-time R01 recipients that acknowledge the R01 grant to all articles we could identify by these authors in *PubMed* within five years of R01 receipt, the overlap exceeded 95%. In other words, first time-R01 recipients' productivity flows almost exclusively into publications emanating from the proposed R01 project. For non-R01 scientists, we considered all

[16] There is no rule that would govern the choice of a ratio for R01s to non-R01s, but fewer controls per R01 scientist produce larger standard errors. Our results remained significant down to an average 1:1 ratio, though, when we matched on deciles rather than quintiles of the matching variables' distributions.

their *PubMed* recorded publications within five years after the applicable observation window.

Dependent variables: We assessed scientific innovation across three domains. First, we calculated innovation productivity as the logged count of publications within five years of the observation window (*ln(publications)*). Scientific publications include innovative findings almost by definition, otherwise they would not merit publication in academic journals[17]. Nonetheless, the degree to which published research represents innovation varies. We therefore used a second measure that captures the research's significance. Parallel to our matching variable, we used the logged count of forward citations per article published within five years of the observation window (*ln(citations)*). To even more directly assess the innovativeness of the research, we also recalculated our novelty matching variable for the five years post observation window (*ln(novelty)*).

4.4 Scientific innovation and funding availability

To identify the marginal impact that more money may have on scientific innovation, we exploited the exogenous shift in NIH funding with a set of indicators that served as our key independent variables.

Independent variables: The first indicator, *budget_period*, equals one if the scientist was observed post the NIH budget doubling decision. This dummy variable captured any change in scientific innovation among non-R01 scientists with the passage of time (counterfactual scientific innovation). The second dummy variable, *R01_grant*, equals one if the scientist received the first R01 grant (pre or post budget doubling), separating the marginal effect of funding from selection into the funding scheme prior to the budget doubling decision. Last, the interaction effect among the two dummy variables, *budget_period X R01_grant*, represented the difference in marginal funding effects on scientific innovation as funding expanded with the NIH budget doubling.

We included additional control variables that are not explicitly featured in the NIH R01 grant evaluation criteria on which we already matched, but that may still influence scientific innovation. Past research has suggested that the timing of the terminal degree may influence academic careers longer term. For example, some cohorts of young scientists may benefit more from generational

[17]Exceptions would be replication studies or reviews, for examples. However, their numbers are small relative to the body of scientists' work and our logged specification should largely nullify their influence.

turnover in academic institutions than others (Blume-Kohout and Clack, 2013). To account for potential cohort effects, we added the first year a scientist appeared on an article in *PubMed* to our regressions (*first_year*). Next, we considered scientist's prior grant experience to potentially influence their ability to produce scientific innovation within a grant cycle of five years, and we included the number of NIH grants received prior to the observation window (*grants_prior*). Since our key independent variables captured the effect of more money on scientific innovation in aggregate (at the federal budget level), we additionally sought to control for funding differences at the individual project level. The *NIH ExPORTER* included annual budget allocations for 1,005 (49%) first-time R01 grants prior to the NIH budget doubling decision and for all first-time R01 grants post the doubling decision. In light of the missing data for our first observation window, we estimated a separate set of regressions that controlled for funding differences at the project level in $100,000 denomination (*R01_money*), disregarding entire groups of matched R01 and non-R01 scientists where we did not have project-level funding information. Note, that we did not have R01 funding information for non-R01 scientists by design[18]. We included fixed effects for the R01 grant vintage (*grant_fe*) in these regressions to account for potential differences on an annual basis beyond the difference from the NIH budget doubling. Finally, we also included fixed effects for area of research inquiry (*field_fe*[19]), and we included fixed effects for each group of matched scientists (*group_fe*).

Table 4 provides descriptive statistics for the matching analysis (upper part) and the scientific innovation analysis (lower part), reported separately for the R01 and matched non-R01 scientists prior to and post the NIH budget doubling decision. The upper part of Table 4 suggests that the matching has effectively selected similar R01 and non-R01 scientists across NIH budget states. The R01 recipients scored somewhat higher than their matches with respect to all our matching variables, irrespective of the NIH budget state. Although not substantive, a certain difference is to be expected if the NIH evaluation process accurately identifies proposals by more competitive scientists for first-time R01 funding. The difference between R01 and non-R01 scientists did not

[18] Although non-R01 scientists are also likely to benefit from some form of funding, the precise level is immaterial for our coefficient estimates as long as we are willing to assume that the amount of funding they receive does not vary substantially across scientists. *Table 5* indicates that over two-thirds of non-R01 scientists operate either under a predoc or postdoc contract, and these salaries are fixed across U.S. institutions.

[19] We used the two-digit letter code for the supporting NIH Institute/Center embedded in the grant number of the last NIH grant received by R01 recipients (i.e., the first R01 grant) and non-R01 scientists (i.e., the last non-R01 NIH grant received).

significantly differ, however, when compared across NIH budget states. Like the results from our logistic regressions of first-time R01 receipt (Table 3), these descriptive statistics indicated a comparable composition of our scientist cohorts across NIH budget states.

The lower part of Table 4 provides descriptive statistics for our scientific innovation models. The difference between R01 and non-R01 scientists substantively widened post the R01 grant receipt across all three measures of scientific innovation. Non-R01 scientists had one to two years less of experience than R01 recipients, based on their first appearance in *PubMed*. The value of R01 grant funding per year increased by about $40,000 (16%)[20]. Non-R01 scientists had on average about half a grant more than first-time R01 recipients. This statistic is indicative of the matched non-R01 scientists being active researchers that appeared to be funded with NIH grants other than the first major R01 grant. Table 5 lists the most recent NIH grant mechanisms that supported our non-R01 scientists. These major mechanisms accounted for more than 80% of the universe of the supporting NIH grants. A minority of non-R01 scientists had only a pre-doctoral level grant on record (e.g., F31). These individuals may, however, have received funding through mechanisms we did not observe (e.g., from institutions other than the NIH) and we kept them for analysis since their publication records matched to a first-time R01 recipient. The majority of non-R01 scientists were funded with post-doctoral (e.g., F32) and career development grants (e.g., K grants). These grants are usually mentored by a more senior scientist and are intended to support the transition to independence. A declining fraction was sponsored by small project grants (about a third of non-R01 scientists) across NIH budget states.

5 Results

Table 6 reports our models assessing the impact of the NIH budget doubling on scientific innovation. We started with a baseline estimation that only included our difference-in-differences dummies and scientist-pair fixed effects that controlled for the characteristics of the R01 recipients and for the variables on which the R01 and non-R01 scientists had been matched. Since all our dependent variables were logged measures of scientific innovation, the exponentiated coefficients on the dummy

[20]This appears in line with broader estimates for how NIH Research Project Grants (RPG) changed in size during the years 1998 to 2003, which have grown by 16%, on average, adjusted by the Biomedical Research and Development Price Index (Scientopia.org). In other words, the change in first-time R01 grant sizes resembles changes in RPG grant sizes more broadly.

variables can be interpreted as percent changes in scientific innovation. Overall, the marginal effect of first-time R01 funding was positive prior to and after the budget doubling decision, though effect sizes varied with the scientific innovation measure considered. Publications within five years of grant award more than doubled for R01 scientists (model 1), prior to ($e^{0.96} = 2.6$) and after the budget expansion ($e^{0.83} = 2.3$). The interaction effect directly estimated the change in the degree to which greater funding with the NIH budget doubling produced scientific innovation (our treatment effect of interest) – publications produced within five years declined by about 12% ($e^{-0.13} = 0.88$). Citations (model 3) and the research's novelty (model 5) also declined with the budget expansion.

These baseline models did not yet account for potential differences by field, which may be important if research areas differed in their budget trajectories during the doubling period. Moreover, the dummy assessing counterfactual scientific innovation (*budget_period*) thus far absorbed both changes across as well as within observation windows (each window spanned three years). We separated the latter from the former by adding scientist cohort fixed effects (*first_year*). We also added the number of prior grants. Overall, including cohort fixed effects increased the dummy on counterfactual publications (model 2) while decreasing it for counterfactual citations (model 4), indicating that non-R01 scientists observed after the budget doubling published more but took a greater hit in their research's significance (citations) when not securing R01 funding as the budget expanded. Field fixed effects did not influence selection across budget phases, suggesting that the allocation of increased funding to first time R01 recipients did not differ meaningfully across research areas (i.e., Institutes)[21]. Looking at the interaction term, these more encompassing model specifications indicated a significant decrease in scientists' innovation productivity (14% fewer papers published), the research's significance (16% less citations generated), and novelty (9% reduction in unprecedented content).

In addition to funding variation at the macro level (federal budget), we also sought to account for variation in funding at the micro level (project budget). Table 7 extends our previous models by adding annual budget dollars (in $100,000 denomination) and R01 grant year fixed effects. Of course, these additions did not change the estimation of changes in scientific innovation for non-R01 scientists (the counterfactual outcome). But it did change the estimation of the treatment effects across measures of scientific innovation. Including actual grant budgets led to a more pronounced

[21] Since field fixed effects were jointly significant, we kept them when estimating additional models.

treatment effect: R01 funding post versus prior to when the NIH budget had doubled yielded 20% fewer publications, 21% less citations, and a 10% reduction in novelty. The inclusion of project budgets changed the treatment effect on publications the most because having more money helped (somewhat) in producing publications (a 6% increase from \$100,000[22]), while it influenced forward citations on these publications less (a 3% improvement), and did nothing to affect the novelty of the research.

5.1 Risk aversity on part of the funder as potential explanation

We predicted that financiers with more money at their disposal pick riskier projects in service of greater scientific innovation. One possible explanation for the negative effects observed may then be, that the NIH did not choose projects for funding that promise greater novelty or significance. Although this may sound somewhat counterintuitive, it is not unfounded. The NIH is generally perceived as a risk averse funding body that, for example, requires substantive preliminary data prior to handing out a grant (Stephan, 2012; Azoulay et al., 2011). The fact that the NIH provides a separate grant mechanism (the R21) to fund the creation of preliminary data for an R01 application may also be viewed as an indication of risk aversity.

To test NIH risk aversity as a possible explanation, we estimated directly how scientific innovation differed with receipt of an R01 grant versus without it, for individuals observed prior to the budget doubling and separately for those observed after the budget doubling. We used our coarsened exact matching (CEM) approach to calculate average treatment effects on the treated (ATET)[23] for each treatment window (1995 to 1997 and 2001 to 2003). To also test whether our results were sensitive to our matching approach, we calculated the ATET using propensity scores and nearest neighbor (NN) matching as an alternative.

Table 8 summarizes the results across the two different matching techniques, across the two time periods (prior and post NIH budget doubling), and provides the corresponding difference-in-differences estimates to facilitate a direct comparison with the results from our pooled models (Table 6). First, the results from the two matching techniques were almost identical for all mea-

[22]This increase would appear modest and echos an analysis conducted by the the National Institute of General Medical Sciences (an Institute of the NIH) that found the bivariate correlation between the number of publications and the total direct cost of grants to be only 0.14 (Berg, 2010).

[23]In contrast to our main models, for these sensitivity analyses we refer to first-time R01 recipients as treated cases since we calculate separate R01 effects for each budget state.

sures of scientific innovation. What is more, the nearest neighbor matching identified a similar number of non-R01 scientists and kept slightly more R01 scientists for analysis. The throughout positive ATETs suggested that the NIH did pick innovative research agendas prior to and after the budget doubling. Perhaps of particular relevance are the positive effects on citations (a 40% to 60% increase with R01 funding) and novelty (a 5% to 14% increase). These findings echo recent research that also provided evidence that the NIH fairly accurately assesses the potential of projects and does pick innovative research for funding (Li and Agha, 2015).

Another explanation for the observed negative effects looks into changes on part of the funded scientists. We consider two potential channels: (i) Allocation of funds lower down the research quality distribution and (ii) scientists changing their research strategy.

5.2 Allocation of funds lower down the quality distribution versus scientists changing their research strategies

Building from the argument and evidence that the NIH appears to be able to assess applicants' proposed research accurately, and given that funds are allocated fairly strictly based on percentile ranks (a 0.1% difference can in theory make or break an application's success), we first entertain the possibility that the additional dollars available from the budget doubling got allocated to research of increasingly lower quality. To build intuition, this scenario would generally translate into a leftward shift of the blue distribution in Figure 1(c). This is because as funding gets extended further down the relative ranking of projects, the quintiles (or any other point on the density function for that matter) should decrease in a lockstep fashion.

If, however, scientists altered their research strategies to pursuing more average (and more certain) projects, the shape of the distribution changes and the shift from the blue to the red distribution in Figure 1(c) should take place. While both channels would lead to a reduction in the mean – explaining our observed average negative treatment effects – the treatment effects should tangibly differ depending on where one looks across the distribution.

We therefore estimate difference-in-differences across distributions' quintiles to probe the relative merit of these two channels. If an expansion of funding led to allocating the marginal dollar to projects of successively lower quality, we would expect similar negative effects of greater funding

availability on scientific innovation across the distributions' quintiles. By contrast, we have argued that lower competition for funding could discourage risk taking, particularly for scientists towards the top end of the distribution who have more to gain when pursuing more average but more certain research strategies that are associated with a high likelihood of funding if about one out of three proposals gets funded[24]. If true, then we would expect a gradient in the treatment effects from top to bottom quintiles. We calculated quintiles from the distributions of our R01 evaluation criteria (matching variables) and placed scientists according to where they fell on the respective distributions at time of potential R01 receipt (our observation windows) and then examined their subsequent research output.

Table 9 compares the combined effect in the bottom four quintiles to the effect in the top quintile of scientists, additionally breaking out the top decile. Going down the second column, we found substantive differences in the estimated treatment effects between the top and bottom quintiles: The data reveal a 1.7x (for publications) to 3x (for novelty) larger reduction in scientific innovation for scientists in the top quintile versus scientists in lower quintiles of the distribution. We even found differences when comparing the top decile to the top quintile. This variation in treatment effects discounted the alternative explanation that the observed decrease in scientific innovation with the expanding NIH budget originated from funding inferior projects across the board[25]. Instead, the evidence appeared most consistent with reduced competition leading particularly top scientists to pursuing more average (and more certain) research strategies relative to scientists in the pre-doubling era, causing a reduction in scientific innovation as funding availability increased.

6 Discussion

A core concern of managers is how firms can grow. Policymakers want the economy as a whole to grow. Scientific innovation spurs economic growth because findings from upstream scientific research often serve as a feedstock for downstream commercial applications. Traditionally, corporations have invested in scientific research (the "R" in R&D) by, for example, hiring researchers

[24] Funding rates for research project grants increased by about one third during the NIH budget doubling to over 35% (Rockey, 2014).

[25] Another piece of evidence that would speak against an allocation of the expanding budget to research of systematically lower quality are the non-significant interaction effects of the R01 evaluation criteria with the NIH budget state dummy in the pooled model of Table 3.

and running basic science labs. Across OECD countries and across industries, corporations have reallocated money away from basic research towards downstream commercialization of research (the "D" in R&D) for the past two decades (Arora et al., 2015). The ensuing corporate research gap is increasingly filled by channeling more money to university research.

The findings from this paper – in the context of the most R&D intensive industry, the life sciences – indicate that increasing the money that flows to university research might reduce scientific innovation across measures of quantity, significance, and novelty. Declining benefits from more money might accumulate over time to eventually stall corporate and economic growth. The top 20 pharmaceutical firms experienced a 464% increase in market capitalization during the 1990–1999 decade, but a 32% decline in market capitalization in the subsequent decade (2000 to 2010). The decline in market capitalization was driven by a precipitous decline in the price-to-earnings multiple whereas earnings based on past innovations increased (Tollman et al., 2011). In other words, investors lost confidence in pharmaceutical firms' ability to develop new products for the future. This episode coincides with the NIH budget doubling, and although the decline in pharmaceutical firms' market capitalization was certainly multifaceted, our findings surface the question to what degree delegating scientific research is a healthy innovation strategy (Freedman et al., 2015), particularly for corporations in innovation-intensive industries.

From a theoretical point of view, our findings therefore suggest value in reinvigorating scholarly attention to the benefits of corporate research, and other corporate activities that bear fruit longer term in the context of highly competitive environments that nurture short-termism. Emerging work quantifies economic growth foregone due to cuts in research expenditures in the range of 0.1% per year (Barton et al., 2017; Barton, 2017; Terry, 2015). Put differently, it seems warranted to examine from a theoretical standpoint under what circumstances firms do stand to gain from investments in research. This could add valuable nuance to the rich literature on the benefits of corporate licensing and joint ventures with universities, and possibly to literature on external knowledge sourcing more broadly (Arora et al., 2001; Arora and Gambardella, 1990).

Another stream of research our findings contribute to is on the science of science and, especially, on incentives for scientific innovation. Money has traditionally taken a back seat relative to incentives in the form of being first to communicate a finding[26], intrinsic satisfaction in solving puzzles,

[26]The importance of this incentive is perhaps best captured by eponymy, the practice of attaching the scientist's

and intellectual autonomy. This literature has usefully documented how these non-monetary incentives spur scientific innovation in the absence of scientists appropriating financial returns from discovery (Stephan, 1996, 2012). It now seems important to complement this view with an examination of how more money may shape both funding allocation and the research strategies of scientists in pursuit of scientific innovation (Franzoni et al., 2011), particularly in heavily funded fields like the life sciences.

Last but not least, our study adds to the broader literature on the economics of innovation. Recent studies that examined the financing of innovation provided grounds to expect a positive relationship between more money being available and innovation outcomes, including in the public funding context (Nanda and Rhodes-Kropf, 2017; Howell, 2017). Innovation starts from an inherently uncertain process that often takes time to yield outcomes. Initial investments generally require more money along the way, giving the financing of innovation an option-like character. A plausible mechanism of how more money promotes innovation then lies in the greater ability to fund higher risk projects because their option value increases with more money being available for additional financing rounds.

To this financier point of view, we add possible motives of the innovator that have remained largely absent from this literature (e.g., Gilbert, 2007)[27]. As more money becomes available for funding, and to the degree that this additional money reduces competition for funding, innovators should be less inclined to pursuing higher risk relative to more certain projects. We find that the NIH budget doubling led to a significant decrease in scientific innovation across the three dimensions of quantity, significance, and novelty. The evidence is consistent with scientists changing their research strategies, particularly at the top end of the distribution. The primary explanation does neither appear to be risk aversity on part of the funder, nor the systematic allocation of funding lower down the quality distribution. The mechanism lies in a reduced incentive to excel, which inheres higher risks, relative to doing better than average.

In addition to the outlined theoretical implications of our results, this study also contributes empirically to the economics of innovation literature in at least two additional respects. There is a

name to a discovery, like Planck's constant or Hodgkin's disease.

[27] A recent exception would be Sauermann and Cohen (2010), who examined how scientists and engineers respond to financial and non-pecuniary rewards, and how that response may affect firm innovation. While this research adds to literature on non-pecuniary rewards in the innovation setting (e.g., Stern, 2004), we were concerned with how money may influence scientists in their decision what research to pursue.

void of modern empirical work on the relationship between science expenditure and output, at least in part attributable to a lack of compelling quasi-experiments and opportunities for observing larger scale changes in research funding more generally (Tabakovic and Wollmann, 2016). Exploiting an exogenous shift in NIH funding, our approach reduces an errors-in-variables problem commonly present in the analysis of longitudinal or cross-sectional data that requires, at a minimum, a large set of control variables for causal estimation that generally also drive effect estimates to zero. Moreover, we observe a substantive shift in funding committed by the world's largest financier of basic research.

An evaluation of the effect of greater R01 funding as a result of the NIH budget expansion might be perceived too narrow still, even if first-time R01 recipients were of similar caliber prior to R01 funding and across NIH budget states. This is because the expanding NIH budget might also affect the research of early career scientists who did not garner first-time R01 funding on the margin, but who might instead receive some bridge funding (e.g., small post-doc grants relative to R01 grants; see Table 5) as more money pours into the entire system. When comparing the effect of greater funding on first-time R01 recipients, one would therefore want to adjust these estimates also with a plausible estimate of research output in the absence of substantial R01 funding across NIH budget states (Jacob and Lefgren, 2011). Overall, it would appear that while the caliber of matched non-R01 scientists was comparable up to the application window for the first R01 grant across NIH budget states (as evidenced in our logistic regressions of the likelihood of R01 receipt and successful matching to eventual first-time R01 scientists thereafter), the subsequent publication records of non-R01 scientists were slightly worse in terms of significance and novelty after the NIH budget expanded (Tables 4 and 6). Said differently, the observed decrease in scientific innovation with more money among first-time R01 recipients were not driven by non-R01 scientists being better after the budget doubling. In fact, we might rather underestimate the decrease in scientific innovation among first-time R01 scientists after the budget expansion because the comparison group not garnering the R01 also performed somewhat worse after the budget expanded.

Like any empirical research, our study involved trade-offs and certain limitations. A more pure empirical estimation in the difference-in-differences logic might have involved simply comparing R01 recipients after to those prior to the budget doubling on grounds of differences in grant packages. While this approach would be feasible in light of the otherwise comparable R01 funding criteria and

pre-publication records of first-time R01 recipients across NIH budget states, it would not account for the differences in the effect of R01 funding isolated from selection effects into the funding scheme. Likewise, one might argue that evaluations of R01 applications are more driven by recent performance (at the extreme by the application itself) than by scientists' publication histories over the past ten years that we use for our matching. Two arguments weighed in favor of our chosen approach. Our matching variables predicted R01 receipt across budget states in a very similar fashion as actual R01 evaluation scores (Eblen et al., 2016). Moreover, if we were to instead build our models from shorter term publication histories, one might surmise that some of the negative effect of more money on scientific innovation might be driven by a regression to the mean effect. In other words, it seems possible that scientists who had recently published highly significant and novel research are more challenged in keeping up the quality, on average. Since we base our matching and our capability-stratified estimates (Table 9) on scientists' longer term publication records, our results seem to more credibly reflect scientists' active decision to pursue less innovative work with R01 funding than scientists being challenged in pursuing innovative work. Lastly, we chose our measures of scientific innovation to be "upstream" (basic science) indicators of innovation to highlight macro implications of channeling more "R&D" *research* money to universities. We do believe, however, that there is value in future work that empirically connects lower scientific innovation springing from university research to possible downstream metrics, like patents and licensing revenues.

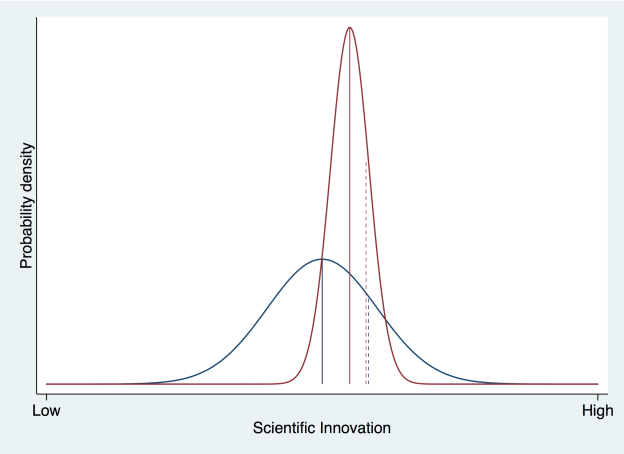
In conclusion, our study provides the first large-scale evidence that allocating more money to university research might not increase but even lessen scientific innovation. If more broadly true, corporations and policymakers may have reason to revisit their innovation strategies. Governments, both in the U.S. and abroad, have come to fund a large share of R&D. Since 2000, the U.S. federal government has accounted for about 30% of total annual U.S. R&D (Boroush, 2016). Perhaps the central assumption in postwar science policy, and increasingly in corporate innovation strategy, is that public funded basic science serves as a broad feedstock for innovative applications (Bush, 1945; Stephan, 2012; Arora et al., 2015). Our findings point at least to a potential opportunity for organizing public funding more productively in service of fostering innovation. We found no tangible upside from larger grant packages after the NIH budget had doubled. Extending funding to more scientists with the use of smaller grants could increase competition (under contracting and

expansionary federal budgets) and may be more effective in stimulating scientific innovation than a universal increase in funding. This perspective would appear relevant on a global scale, observing the recent injections of public research funding in Canada, China, and Japan, for examples. More generally, our study points to the importance of revisiting the division of labor in R&D and the allocation of resources in support of scientific innovation.

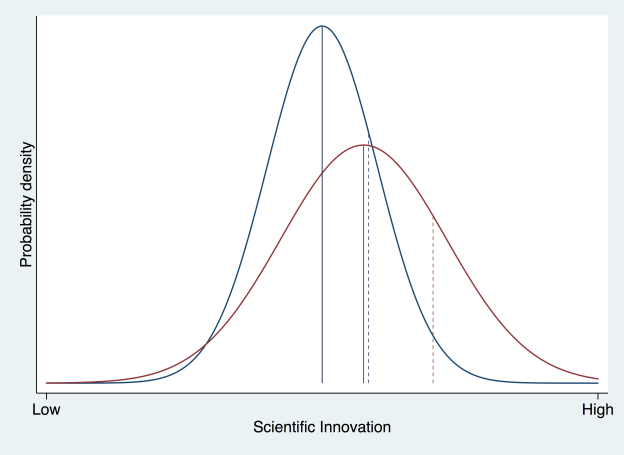
7 Appendix I: Scientist Disambiguation

Since scientists' publication histories were crucial to our research design, we needed to ensure that two or more instances of the same name (or of highly similar names) on different papers actually represented the same NIH funded scientist we cared about. The *Author-ity* database uses an algorithm that incorporates information on shared title words, journal names, coauthors, medical subject headings, publication language, affiliations, email addresses, and author name features (middle initial, suffix, and name prevalence) to determine unique author IDs for *PubMed* (Torvik and Smalheiser, 2009). Meanwhile, the *NIH ExPORTER* database contains two unique identifiers – one for each awarded grant and one for each principal investigator (PI ID) funded by the NIH. Of note, the PI IDs remain constant from project to project and from year to year. Since the *NIH ExPORTER* also connects publications that emanate from grants, recording the same unique identifier for articles as used in the *PubMed* database, we can use these unique article identifiers (pmids) as a crosswalk between the NIH PI IDs and the *Author-ity* author IDs. Prior work has documented that individuals' publication histories can be assembled from these three data sources with over 99% accuracy, over time and across different levels of scientists' productivity and name prevalence (Lerchenmueller and Sorenson, 2016). Importantly, this author disambiguation allowed us to not only assemble the post-grant publication histories of NIH funded scientists but also their pre-grant publication histories that we use to compare the caliber of first-time R01 recipients as well as create our set of scientists that would appear similarly eligible for first-time R01 funding but who did not receive the R01 grant yet.

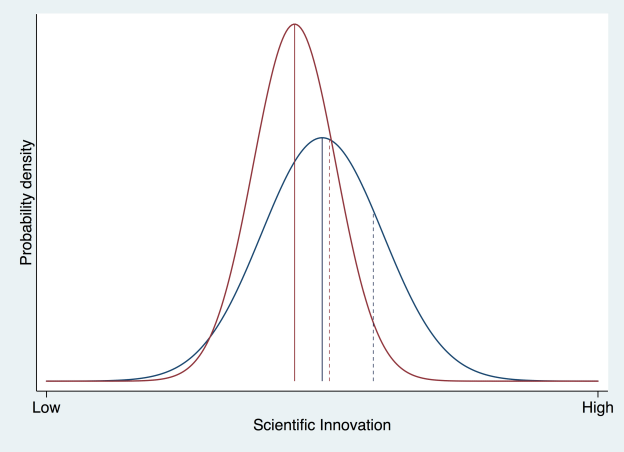
Figure 1: Potential mechanisms relating greater availability of funding and scientific innovation



(a) Money funds better resources



(b) Financiers pick riskier projects



(c) Scientists pursue more average projects

Note: Stylized probability density functions; the shift from blue to red distribution varies with mechanism considered; solid vertical and dashed lines mark the mean and 80^{th} percentile of the respective distributions.

Figure 2: First time R01 grants and associated research outcomes within five years of grant award

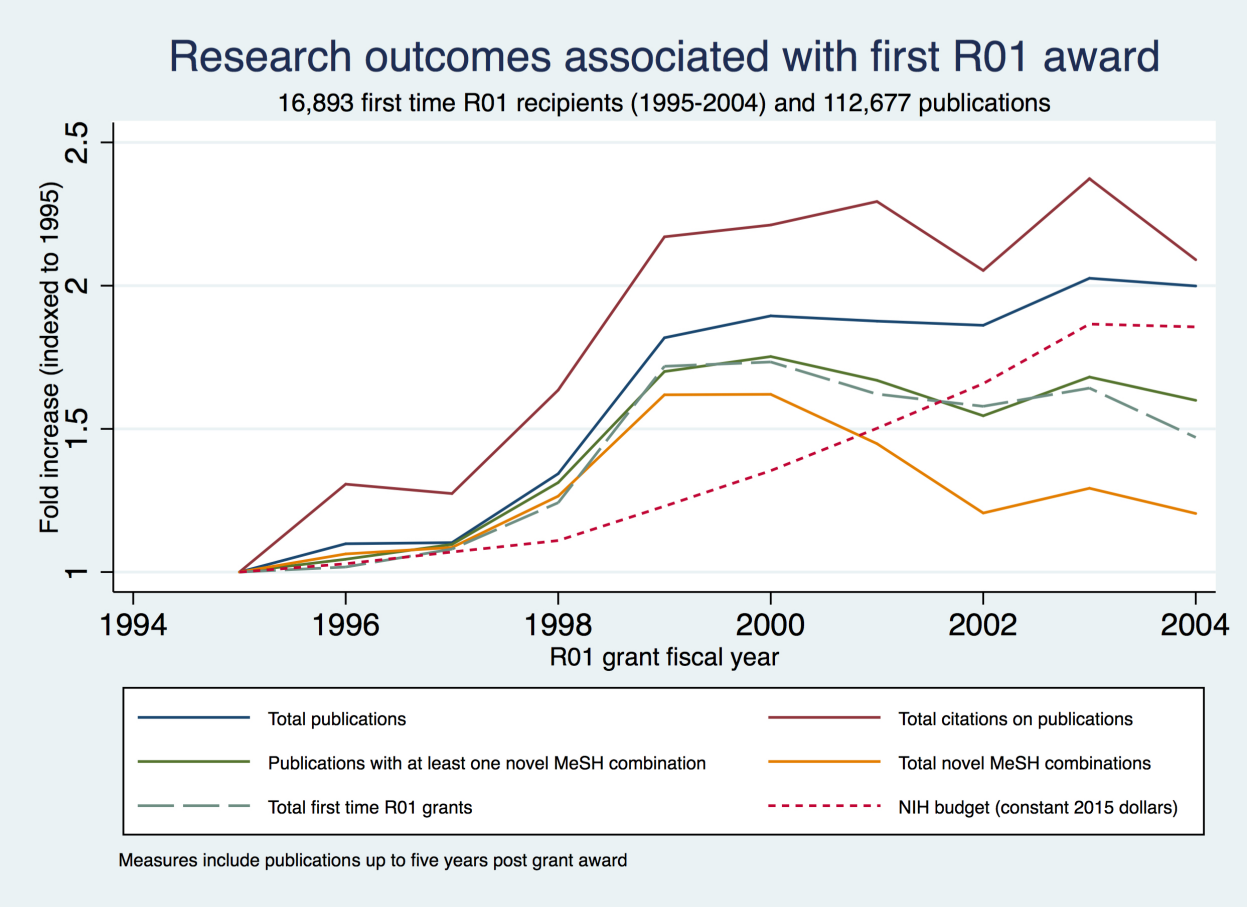


Table 1: Definition of Variables

Variable	Description	Type*
Ln(pub_prior)	ln of total number of articles published up to 10 years prior to observation window, adjusted for years of research experience	MV
Ln(cites_prior)	ln of forward citations per article published up to 10 years prior to observation window	MV
Ln(novelty_prior)	ln of the average proportion of new keyword (MeSH term) combinations across impact-weighted articles published up to 10 years prior to observation window	MV
Org_status	Host institution's percentile rank in terms of nationwide R01 funding during observation window (0 to 1)	MV
Sex	Dummy variable equal to 1 if scientist is female	MV
Ln(publications)	ln of total number of articles published within five years of observation window	DV
Ln(citations)	ln of forward citations per article published within five years of observation window	DV
Ln(novelty)	ln of the average proportion of new keyword (MeSH term) combinations across impact-weighted articles published within five years of observation window	DV
Budget_period	Dummy variable equal to 1 if scientist belongs to cohort post NIH budget doubling decision	IV
R01_grant	Dummy variable equal to 1 if scientist received the first R01 grant prior to or post the NIH budget doubling decision	IV
Budget_period X R01_grant	Dummy variable equal to 1 if scientist received R01 grant and R01 grant receipt occurred between 2001 and 2003 (i.e., after the NIH budget doubling decision)	IV
First_year	Year of scientist's first appearance as an author in *PubMed*	CV
Grants_prior	Scientist's total NIH grants prior to observation window	CV
R01_money	R01 grant package, expressed in $100,000 per average grant year	CV
Field_FE	Research field fixed effects (21 dummy variables, one for each funding NIH Institute)	CV
Group_FE	Scientist pair fixed effects ($> 1,000$ dummy variables, one for each group of matched scientists)	CV
Grant_FE	R01 grant vintage fixed effects (six dummy variables, one for each observation window year)	CV

*MV (matching variable), DV (dependent variable), IV (independent variable), CV (additional control variable)

Table 2: Evaluation Criteria for NIH R01 Grant Applications and Matching Variables

Criteria*	Description	Matching variable
1. Significance	Will scientific knowledge be advanced by completing the proposed research ?	*Ln(cites_prior)*
2. Investigator	Does PI have appropriate experience and an ongoing record of accomplishments ?	*Ln(pub_prior)*
3. Innovation	Does proposal shift current research with new concepts and methodologies ?	*Ln(novelty_prior)*
4. Approach	Are strategy and methods appropriate to accomplish the proposed research ?	*Ln(pub_prior)*
5. Environment	Will the scientific environment contribute to probability of success ?	*Org_status*

*To address potential (un)conscious bias we additionally match on scientist's sex

Source: NIH R01 evaluation criteria

Table 3: Logistic regression of likelihood of receiving first R01 grant in dependence of review criteria

	Odds Ratio	Robust S.E.	95% CI lower	95% CI upper
1. Prior to doubling				
Ln(pub_prior)	4.31	0.31	3.75	4.96
Ln(cites_prior)	1.58	0.06	1.47	1.70
Ln(novelty_prior)	1.23	0.06	1.11	1.36
Org_status	1.02	0.00	1.01	1.02
Sex	0.83	0.05	0.74	0.93
2. Post doubling				
Ln(pub_prior)	3.80	0.22	3.39	4.27
Ln(cites_prior)	1.69	0.05	1.59	1.80
Ln(novelty_prior)	1.29	0.06	1.18	1.41
Org_status	1.02	0.00	1.01	1.02
Sex	0.85	0.04	0.77	0.94
3. Pooled model				
Ln(pub_prior)	4.30	0.30	3.75	4.93
Ln(cites_prior)	1.55	0.06	1.45	1.66
Ln(novelty_prior)	1.21	0.06	1.09	1.33
Org_status	1.02	0.00	1.01	1.02
Sex	0.84	0.04	0.76	0.93
Post_budget	0.97	0.27	0.56	1.69
Post_budget X				
Ln(pub_prior)	0.88	0.08	0.74	1.06
Ln(cites_prior)	1.09	0.05	0.99	1.20
Ln(novelty_prior)	1.09	0.07	0.95	1.24
Org_status	1.00	0.00	1.00	1.00
Sex	1.00	0.08	0.86	1.16

Note: All matching variables predict R01 ($p < 0.05$); post budget dummy and interactions not significant. Model (1) uses 7,780 and model (2) uses 10,063 and the pooled model (3) uses 17,843 observations.

Table 4: Summary Statistics of Matched R01 and Non-R01 Scientists Prior and Post NIH Budget Doubling

	R01s prior		Non-R01s prior		R01s post		Non-R01s post	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Matching								
Ln(pub_prior)	1.11	0.43	0.91	0.41	1.11	0.41	0.93	0.42
Ln(cites_prior)	2.25	0.96	1.85	0.96	2.57	0.87	2.19	0.91
Ln(novelty_prior)	-1.25	0.69	-1.48	0.63	-1.37	0.58	-1.54	0.58
Org_status	95.00	9.79	89.38	23.36	94.61	9.69	87.35	26.12
Sex (pct. women)	27.65	44.74	31.28	46.37	33.93	47.35	39.95	48.98
Innovation								
Ln(publications)	2.89	0.74	1.65	1.04	2.93	0.74	1.82	1.07
Ln(citations)	2.41	0.81	1.61	1.11	2.41	0.69	1.78	1.07
Ln(novelty)	-1.42	0.53	-1.62	0.58	-1.60	0.45	-1.72	0.52
Controls								
First_year	1986.4	4.19	1988.0	4.61	1993.4	3.88	1995.6	4.63
Grants_prior	0.77	0.92	1.24	0.62	0.91	1.02	1.27	0.84
R01_money (in 100,000)	264.3*	134.1*			307.5	121.0		
N	2,051		4,098		2,868		5,530	

*Based on 1,005 first-time R01 recipients whose grants contained dollar funding information in the NIH ExPORTER.

Table 5: NIH Grant Mechanisms Supporting Non-R01 Scientists

Prior and Post the NIH Budget Doubling

Grant type	Share (%) non-R01s prior	Share (%) non-R01s post
Pre-doctoral grants (e.g., F31)	11%	12%
Post-doctoral grants (e.g., F32, K01, K08)	51%	60%
Small research grants (e.g., R03, R21, R29)	38%	28%

Pre- and post-doc grants are individual career development grants, often with mentorship.

Small research grants are, by contrast, mostly for supporting the execution of a specific project.

Note: The NIH discontinued the R29 award in 1998, redirecting funding to the R21 mechanism.

Table 6: Difference-in-differences estimation of average treatment effect of greater funding availability on scientific innovation

	Ln(publications)		*Ln(citations)*		*Ln(novelty)*	
	(1)	(2)	(3)	(4)	(5)	(6)
Main effects						
Budget_period	0.24***	0.45***	0.20***	0.14***	-0.09**	-0.08**
	(0.03)	(0.04)	(0.03)	(0.04)	(0.02)	(0.02)
R01_grant	0.96***	0.94***	0.50***	0.51***	0.12***	0.13***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
Budget_period X	-0.13***	-0.15***	-0.18***	-0.17***	-0.08***	-0.09**
R01_grant	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
Controls						
First_year		-0.03***		0.01**		-0.00
		(0.00)		(0.00)		(0.00)
Grants_prior		0.03**		0.00		0.00
		(0.01)		(0.01)		(0.01)
Group_FE (1,063)	YES	YES	YES	YES	YES	YES
Field_FE (20)	NO	YES	NO	YES	NO	YES
R^2	0.46	0.48	0.34	0.36	0.24	0.27
F	721.91	116.33	164.37	29.41	30.07	19.23
Observations	14,547	14,547	14,574	14,574	14,574	14,574

Significance levels: † 10%; * 5%; ** 1%; *** 0.1%; clustered standard errors in parentheses.

Table 7: Difference-in-differences estimation of average treatment effect of greater funding availability on scientific innovation with project budget fixed effects

	Ln(publications)	*Ln(citations)*	*Ln(novelty)*
Main effects			
Budget_period	0.44***	0.13**	-0.08***
	(0.04)	(0.04)	(0.02)
R01_grant	0.67***	0.47***	0.13***
	(0.05)	(0.05)	(0.03)
Budget_period X	-0.23***	-0.24***	-0.11***
R01_grant	(0.05)	(0.05)	(0.03)
Controls			
First_year	-0.03***	0.01**	-0.00
	(0.00)	(0.00)	(0.00)
Grants_prior	0.03*	0.00	-0.00
	(0.01)	(0.01)	(0.01)
R01_money	0.06***	0.03*	0.00
	(0.01)	(0.01)	(0.01)
Group_FE (1,063)	YES	YES	YES
Field_FE (20)	YES	YES	YES
Grant_FE (5)	YES	YES	YES
R^2	0.48	0.36	0.26
F	107.73	24.06	14.72
Observations	13,501	13,501	13,501

Significance levels: † 10%; * 5%; ** 1%; *** 0.1%; clustered standard errors in parentheses.

Table 8: Average treatment effect on the treated (ATET) prior and post NIH budget doubling with nearest neighbor (NN) matching vs. coarsened exact matching (CEM)

	NN		CEM		
	Prior	Post	Prior	Post	Δ
Ln(publications)					
ATET	0.96	0.83	0.93	0.78	
diff-in-diff estimate					-0.15
Ln(citations)					
ATET	0.50	0.32	0.50	0.34	
diff-in-diff estimate					-0.17
Ln(novelty)					
ATET	0.12	0.04	0.13	0.05	
diff-in-diff estimate					-0.09
N cases	2,113	2,910	2,051	2,868	
N controls	4,121	5,536	4,098	5,530	
Total N (CEM)					14,547

Note: All average treatment effects on the treated are significant at the 1% level.

Table 9: Difference-in-differences estimation by top decile and quintile versus combined bottom four quintiles of scientists' capability distribution

	Estimate	S.E.	Lower CI	Upper CI
Ln(publications)				
Top decile	-0.19	0.08	-0.34	-0.03
Top quintile	-0.22	0.06	-0.35	-0.10
Bottom quintiles	-0.13	0.04	-0.20	-0.05
Ln(citations)				
Top decile	-0.29	0.11	-0.50	-0.08
Top quintile	-0.26	0.07	-0.40	-0.13
Bottom quintiles	-0.13	0.04	-0.20	-0.06
Ln(novelty)				
Top decile	-0.18	0.11	-0.40	0.03
Top quintile	-0.16	0.05	-0.27	-0.06
Bottom quintiles	-0.05	0.02	-0.10	-0.01

Note: Standard errors (S.E.) are clustered at the scientist-pair level; 95% confidence bounds shown.

References

Arora, Ashish, Sharon Belenzon, Andrea Patacconi. 2015. Killing the golden goose? The decline of science in corporate R&D. Tech. rep., National Bureau of Economic Research.

Arora, Ashish, Andrea Fosfuri, Alfonso Gambardella. 2001. *Markets for technology: The economics of innovation and corporate strategy*. MIT Press, Cambridge.

Arora, Ashish, Alfonso Gambardella. 1990. Complementarity and external linkages: The strategies of the large firms in biotechnology. *The Journal of Industrial Economics* 361–379.

Audretsch, David B., Maryann P. Feldman. 1996. R&D spillovers and the geography of innovation and production. *The American Economic Review* **86**(3) 630–640. URL `http://www.jstor.org/stable/2118216`.

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

Barton, Dominic. 2017. Refocusing capitalism on the long term: Ownership and trust across the investment value chain. *Oxford Review of Economic Policy* **33**(2) 188–200. doi:10.1093/oxrep/grx025. URL `+http://dx.doi.org/10.1093/oxrep/grx025`.

Barton, Dominic, James Manyika, Sarah Keohane Williamson. 2017. Finally, evidence that managing for the long term pays off. *Harvard Business Review* URL `https://hbr.org/2017/02/finally-proof-that-managing-for-the-long-term-pays-off`.

Basalla, G. 1988. *The Evolution of Technology*. Cambridge University Press, Cambridge, MA.

Becker, Gary S. 1994. Human capital revisited. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*. The University of Chicago Press, 15–28.

Benavente, José Miguel, Gustavo Crespi, Lucas Figal Garone, Alessandro Maffioli. 2012. The impact of national research funds: A regression discontinuity approach to the Chilean FONDECYT. *Research Policy* **41**(8) 1461–1475.

Berg, J. 2010. Another look at measuring the scientific output and impact of NIGMS grants. Tech. rep., NIGMS Feedback Loop. URL `https://loop.nigms.nih.gov/2010/11/another-look-at-measuring-the-scientific-output-and-impact-of-nigms-grants/`.

Bikard, Michaël, Fiona Murray, Joshua S. Gans. 2015. Exploring trade-offs in the organization of scientific work: Collaboration and scientific reward. *Management Science* **61**(7) 1473–1495. doi:10.1287/mnsc.2014.2052. URL `https://doi.org/10.1287/mnsc.2014.2052`.

Blume-Kohout, Margaret E., John W. Clack. 2013. Are graduate students rational? Evidence from the market for biomedical scientists. *PLOS ONE* **8**(12) 1–13. doi:10.1371/journal.pone.0082759. URL https://doi.org/10.1371/journal.pone.0082759.

Boroush, Mark. 2016. U.S. R&D increased by more than $20 billion in both 2013 and 2014, with similar increase estimated for 2015. Tech. rep., National Science Foundation. URL https://www.nsf.gov/statistics/2016/nsf16316/#fn1.

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, Christoph Riedl. 2016. Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science* **62**(10) 2765–2783. doi:10.1287/mnsc.2015.2285. URL https://doi.org/10.1287/mnsc.2015.2285.

Bush, V. 1945. Science: The endless frontier. Tech. rep., U.S. General Printing Office, Washington, DC.

DeLong, Bradford J. 2000. Cornucopia: The pace of economic growth in the twentieth century. *NBER Working Paper 7602. National Bureau of Economic Research, Cambridge, MA* .

Dewar, Robert D., Jane E. Dutton. 1986. The adoption of radical and incremental innovations: An empirical analysis. *Management Science* **32**(11) 1422–1433. doi:10.1287/mnsc.32.11.1422. URL https://doi.org/10.1287/mnsc.32.11.1422.

Eblen, Matthew K, Robin M Wagner, Deepshikha RoyChowdhury, Katherine C Patel, Katrina Pearson. 2016. How criterion scores predict the overall impact score and funding outcomes for National Institutes of Health peer-reviewed applications. *PloS ONE* **11**(6) e0155060.

Ederer, Florian, Gustavo Manso. 2013. Is pay for performance detrimental to innovation? *Management Science* **59**(7) 1496–1513. doi:10.1287/mnsc.1120.1683. URL https://doi.org/10.1287/mnsc.1120.1683.

Eisenhardt, Kathleen M. 1989. Agency theory: An assessment and review. *Academy of Management Review* **14**(1) 57–74.

Fleming, Lee. 2001. Recombinant uncertainty in technological search. *Management Science* **47**(1) 117–132. doi:10.1287/mnsc.47.1.117.10671. URL https://doi.org/10.1287/mnsc.47.1.117.10671.

Fleming, Lee. 2007. Breakthroughs and the "long tail" of innovation. *MIT Sloan Management Review* **49**(1) 69–74.

Fleming, Lee, Olav Sorenson. 2004. Science as a map in technological search. *Strategic Management Journal* **25**(8-9) 909–928. doi:10.1002/smj.384. URL http://dx.doi.org/10.1002/smj.384.

Franzoni, Chiara, Giuseppe Scellato, Paula Stephan. 2011. Changing incentives to publish. *Science* **333**(6043) 702–703.

Freedman, Leonard P., Iain M. Cockburn, Timothy S. Simcoe. 2015. The economics of reproducibility in preclinical research. *PLOS Biology* **13**(6) 1–9. doi:10.1371/journal.pbio.1002165. URL `https://doi.org/10.1371/journal.pbio.1002165`.

Gilbert, Richard J. 2007. Competition and innovation. Tech. rep., UC Berkeley: Competition Policy Center. Retrieved from https://escholarship.org/uc/item/9xh5p5p9.

Gilfillan, S. C. 1935. *The Sociology of Invention*. Follett Publishing, Chicago.

Hall, Bronwyn, Josh Lerner. 2010. The financing of R&D and innovation. Bronwyn Hall, Nathan Rosenberg, eds., *The Handbook of the Economics of Innovation*, vol. 1, chap. 2. Elsevier Press, London MA, 609–639.

Hargadon, Andrew, Robert I. Sutton. 1997. Technology brokering and innovation in a product development firm. *Administrative Science Quarterly* **42**(4) 716–749.

Howell, Sabrina T. 2017. Financing innovation: Evidence from R&D grants. *The American Economic Review* **107**(4) 1136–1164.

Iacus, Stefano M, Gary King, Giuseppe Porro. 2012. Causal inference without balance checking: Coarsened exact matching. *Political Analysis* **20**(1) 1–24.

Jacob, Brian A, Lars Lefgren. 2011. The impact of research grant funding on scientific productivity. *Journal of Public Economics* **95**(9-10) 1168–1177.

Janis, I. 1972. *Victims of Groupthink*. Houghton Mifflin Company, Boston.

Johnson, Judith A. 2016. NIH Funding: FY1994-FY2016. Tech. rep., Congressional Research Service, Washington, DC.

Jones, Benjamin F. 2009. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies* **76**(1) 283–317.

Jones, Benjamin F., Stefan Wuchty, Brian Uzzi. 2008. Multi-university research teams: Shifting impact, geography, and stratification in science. *Science* **322**(5905) 1259–1262. doi:10.1126/science.1158357. URL `http://science.sciencemag.org/content/322/5905/1259`.

Klingebiel, Ronald, Christian Rammer. 2014. Resource allocation strategy for innovation portfolio management. *Strategic Management Journal* **35**(2) 246–268.

Kogut, Bruce, Udo Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* **3**(3) 383–397. doi:10.1287/orsc.3.3.383. URL `https://doi.org/10.1287/orsc.3.3.383`.

Lauer, Mike. 2014. Are attempts at renewal successful? Tech. rep., National Institutes of Health. URL `https://nexus.od.nih.gov/all/2016/02/16/are-attempts-at-renewal-successful/`.

Lerchenmueller, Marc, 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.

Lerchenmueller, Marc, Olav Sorenson. 2018. The gender gap in early career transitions in the life sciences. *Research Policy* Forthcoming.

Li, Danielle, Leila Agha. 2015. Big names or big ideas: Do peer-review panels select the best science proposals? *Science* **348**(6233) 434–438. doi:10.1126/science.aaa0185. URL `http://science.sciencemag.org/content/348/6233/434`.

Li, Danielle, Pierre Azoulay, Bhaven N. Sampat. 2017. The applied value of public investments in biomedical research. *Science* **356**(6333) 78–81. doi:10.1126/science.aal0010. URL `http://science.sciencemag.org/content/356/6333/78`.

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

March, James G. 1991. Exploration and exploitation in organizational learning. *Organization Science* **2**(1) 71–87. doi:10.1287/orsc.2.1.71. URL `https://doi.org/10.1287/orsc.2.1.71`.

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

Mokyr, Joel. 2010. The contribution of economic history to the study of innovation and technical change: 1750–1914. Bronwyn Hall, Nathan Rosenberg, eds., *The Handbook of the Economics of Innovation*, vol. 1, chap. 2. Elsevier Press, London MA, 11–50.

Moses, H, Matheson DM, Cairns-Smith S, George BP, Palisch C, Dorsey E. 2015. The anatomy of medical research: US and international comparisons. *JAMA* **313**(2) 174–189. doi:10.1001/jama.2014.15939. URL `http://dx.doi.org/10.1001/jama.2014.15939`.

Nanda, Ramana, Matthew Rhodes-Kropf. 2017. Financing risk and innovation. *Management Science* **63**(4) 901–918. doi:10.1287/mnsc.2015.2350. URL `https://doi.org/10.1287/mnsc.2015.2350`.

National Library of Medicine. 2017. PubMed Database. Data available via online interface at, `https://www.ncbi.nlm.nih.gov/pubmed/`.

NIH. 2017a. Next generation researchers initiative. Tech. rep., National Institutes of Health. URL `https://grants.nih.gov/ngri.htm`.

NIH. 2017b. NIH ExPORTER. Data retrieved from, `https://exporter.nih.gov/default.aspx`.

Nonaka, Ikujiro. 1994. A dynamic theory of organizational knowledge creation. *Organization Science* **5**(1) 14–37.

Rockey, Sally. 2014. Comparing success rates, award rates, and funding rates. Tech. rep., National Institutes of Health. URL https://nexus.od.nih.gov/all/2014/03/05/comparing-success-award-funding-rates/.

Rothaermel, Frank T., Andrew M. Hess. 2007. Building dynamic capabilities: Innovation driven by individual-, firm-, and network-level effects. *Organization Science* **18**(6) 898–921. doi:10.1287/orsc.1070.0291. URL https://doi.org/10.1287/orsc.1070.0291.

Sachs, Frederick. 2007. Is the NIH budget saturated? Tech. rep., The Scientist. URL http://www.the-scientist.com/?articles.view/articleNo/25416/title/Is-the-NIH-budget-saturated-/.

Sampat, Bhaven N, Frank R Lichtenberg. 2011. What are the respective roles of the public and private sectors in pharmaceutical innovation? *Health Affairs* **30**(2) 332–339.

Sauermann, Henry, Wesley M. Cohen. 2010. What makes them tick? Employee motives and firm innovation. *Management Science* **56**(12) 2134–2153. doi:10.1287/mnsc.1100.1241. URL https://doi.org/10.1287/mnsc.1100.1241.

Schumpeter, J. 1939. *Business Cycles*. McGraw-Hill: New York.

Singh, Jasjit, Lee Fleming. 2010. Lone inventors as sources of breakthroughs: Myth or reality? *Management Science* **56**(1) 41–56. doi:10.1287/mnsc.1090.1072. URL https://doi.org/10.1287/mnsc.1090.1072.

Stephan, Paula E. 1996. The economics of science. *Journal of Economic Literature* **34**(3) 1199–1235. URL http://www.jstor.org/stable/2729500.

Stephan, Paula E. 2012. *How Economics Shapes Science*, vol. 1. Harvard University Press Cambridge, MA.

Stern, Scott. 2004. Do scientists pay to be scientists? *Management Science* **50**(6) 835–853. doi:10.1287/mnsc.1040.0241. URL https://doi.org/10.1287/mnsc.1040.0241.

Tabakovic, Haris, Thomas Wollmann. 2016. The impact of money on science: Evidence from unexpected NCAA football outcomes. Working Paper, Chicago Business School.

Terry, Stephen J. 2015. The macro impact of short termism. URL http://economics.mit.edu/files/10386. Working Paper, Boston University.

Tollman, P, Y Morieux, J Murphy, U Schulze. 2011. Can R&D be fixed? Lessons from Biopharma outliers. *BCG Perspectives* .

Torvik, Vetle I., Neil R. Smalheiser. 2009. Author name disambiguation in MEDLINE. *ACM Trans. Knowl. Discov. Data* **3**(3) 11:1–11:29. URL http://doi.acm.org/10.1145/1552303.1552304.

Weitzman, Martin L. 1998. Recombinant growth. *The Quarterly Journal of Economics* **113**(2) 331–360.

Whalley, Alexander, Justin Hicks. 2014. Spending wisely? How resources affect knowledge production in universities. *Economic Inquiry* **52**(1) 35–55.

Wuchty, Stefan, Benjamin F. Jones, Brian Uzzi. 2007. The increasing dominance of teams in production of knowledge. *Science* **316**(5827) 1036–1039. doi:10.1126/science.1136099. URL `http://science.sciencemag.org/content/316/5827/1036`.

Zerhouni, Elias A. 2006. NIH in the post-doubling era: Realities and strategies. *Science* **314**(5802) 1088–1090.

Zucker, Lynne G, Michael R Darby. 1996. Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences* **93**(23) 12709–12716.