

Scientific Grant Funding

Pierre Azoulay and Danielle Li*
MIT & NBER

March 4, 2020

Abstract

This chapter provides an overview of grant funding as an innovation policy tool aimed at both practitioners and science policy scholars. We first discuss how grants relate to other contractual mechanisms such as patents, prizes, or procurement contracts, and argue that, among these, grants are likely to be the most effective way of supporting early stage, exploratory science. Next, we provide a brief history of the modern scientific grant and discuss the current state of knowledge regarding several key elements of the design of grant programs: the choice of program scope, the design of peer review, as well as approaches for creating incentives for risk-taking and translation for grant recipients. We argue that, in making these choices, policy-makers might consider adopting a portfolio-based mindset that seeks a diversity of approaches, while accepting that high failure rates for individual projects is in fact part of an effective grant-making program. Finally, increased rigor in the evaluation of grant programs is likely to raise the quality of funded proposals. In particular, randomized controlled trials and other quasi-experimental techniques might enable policy makers to communicate and enhance the impact that these programs have on discovery and innovation, thereby creating a stronger justification for their expansion or continued existence.

The pharmaceutical firm Novartis made use of decades of research in the development of Gleevec, a remarkably effective treatment for chronic myelogenous leukemia (CML). Between the 1960s and 1980s, numerous studies funded by the National Institutes of Health (NIH) investigated the causes of CML, documenting the role of a specific gene mutation that leads tyrosine kinase, a common cell signaling molecule, to become overactive. This understanding pointed to an approach for treating CML—develop compounds to inhibit tyrosine kinase—which Novartis scientists then pursued. Beyond treating CML, Gleevec also served as a proof-of-concept that ushered in a new era of targeted cancer therapeutics (Wapner 2013).

Similarly, the National Science Foundation did not anticipate laying the foundation for secure internet commerce when it awarded grant MCS76-74294 in to a young MIT assistant professor Ronald Rivest with the general-purpose title “concrete computational complexity.” Yet Rivest, together with colleagues Adi Shamir and Leonard Adleman used these funds to develop the first public-key cryptosystem, thus revolutionizing the field of cryptography and enabling a myriad of applications for the transmission of data using digital signatures (Rivest et al. 2018).

* Send correspondence to pazoulay@mit.edu. We thank Ben Jones and Austan Goolsbee for useful suggestions.

While impacting different sectors of the economy, Gleevec and the RSA algorithm are innovations that share three essential traits. First, although they were eventually commercialized by private firms, each owes a clear intellectual debt to research grants awarded by public sector entities: the NIH, DOD, and NSF. Second, these grant funds were not earmarked with these specific outcomes in mind, but rather were given for general inquiries in the fields of genetics and theoretical computer science, without any conditions with respect to the purported “usefulness” of the recipients’ work. Lastly, while these projects eventually lead to tremendous societal gains, many other projects supported by these same agencies either failed outright or only generated incremental benefits.

These features capture both the promise and pitfalls of investing in basic science: while nascent ideas have the potential to have widespread and substantial impacts, it is very difficult to predict whether, when, or how. Moreover, even when the value of investments is clear, as in the cases above, it is often difficult to quantify. Together, this lack of predictability and traceability has made grant funding politically vulnerable.

Emerging research, however, has begun to provide concrete evidence that grants play a critical role in enabling and sustaining innovation. In their studies of funding for biomedical research, for instance, Li et al. (2017) and Azoulay et al. (2019a) show that NIH funded research lays the foundation upon which private sector science builds: over 40% of NIH-funded grants produce research that is cited by a private sector patent, and a single dollar in NIH funding translates into private sector spillovers worth twice that amount, not counting any direct value of academic research or training. Howell (2017) studies applicants to the US Department of Energy’s Small Business Innovation Research (SBIR) grant program, and finds that early-stage awards approximately double the probability that a firm receives subsequent venture capital, and has a large and positive impact on patenting and revenues. Her results are consistent with the view that non-dilutive funding of this type allows small firms to fund technology prototyping, thereby accelerating the translation of academic results into useful products.

Economists and historians have long acknowledged the key role played by institutions in translating scientific knowledge into welfare-enhancing innovations (Dasgupta and David 1994; Mokyr 2002; Rosenberg 1979). Perhaps because grant systems are ubiquitous in the research world, they have been treated as an immutable, taken-for-granted background institution for financing basic research. Relative to prizes or patents, they have received less scholarly attention, and the ample theoretical literature on procurement (Laffont and Tirole 1993) does not appear to recognize grants as a distinct class of contractual devices, offering at best a very stylized treatment of their optimal use and design (Wright 1983; Gallini and Scotchmer 2002).¹ Yet, in a growing acknowledgement of its importance, empirical studies in the past decade have begun to examine the relationship between specific modes of science funding and the rate and direction of scientific inquiry. The present chapter reviews the literature on scientific grants, in an effort to

¹ A notable exception is the work of Price (2019), who provides a comprehensive review of grant funding from a legal scholar’s perspective.

suggest promising avenues for reforming this important and necessary—but under-studied—contractual mechanism.

Throughout, we emphasize three themes.

First, grants, patents, prizes, and research contracts play overlapping and mutually supportive roles in the research funding ecosystem, with grants most effective when research is exploratory, and when it is likely to produce ample spillovers, both across domains and over time. These two features characterize much early-stage scientific research.

Second, grant programs must be designed in ways that recognize the possibility of failure. This entails encouraging recipients to take on scientific and technological risks, exploring new research avenues rather than sticking with safer and more conventional trajectories.

Third, funding agencies could consider encouraging the systematic evaluation of grant programs by comparing outcomes amongst scientists, institutions, or fields that receive funding with those that accrue to “control” scientists, institution, or fields that do not.

The chapter proceeds as follows. In the next section, we identify the circumstances under which grants may be preferred over alternatives such as patents, prizes, or traditional procurement contracts. After providing a brief history of scientific grantmaking, we highlight key design choices faced by science policy makers when setting up a grant system: (1) delineating the scope of the grant competition and the set of potential applicants; (2) choosing a method to select meritorious applications; (3) providing incentives for the winning applicants; and (4) evaluating outcomes. We discuss the current state of knowledge regarding trade-offs entailed by alternative design choices in each of these domains, highlighting many questions still open in light of the extant evidence. We conclude with a robust defense of scientific grants as a necessary tool in the wider ecosystem of R&D funding, and suggest that funders consider using randomized, controlled experimentation as a way of identifying the specific funding practices worthy of systematic adoption—or abandonment.

1. Why fund scientific research through grants?

Ever since Vannevar Bush’s report *Science: The Endless Frontier*, US policy makers have generally agreed that basic scientific research “creates the fund of new knowledge from which the practical applications of knowledge must be drawn” (Bush 1945). Because scientific knowledge often exhibits the characteristics of a public good, economists have argued that it would be underprovided by the private sector (Nelson 1959,

Arrow 1962), thereby providing a rationale for public expenditures devoted to the funding of scientific research.²

However, it is one thing to argue that there is role for public support of scientific investments, and quite another to determine what specific form this support should take. In this chapter, we examine one specific set of institutions often employed in advanced economies to fund investments in research: scientific grant funding.

Grants are upfront payments for the delivery of incompletely specified and non-contractable R&D output. Unlike research prizes, the funder must pay before there is any guarantee of a successful outcome. Unlike loans, one cannot ask for grant money back if a project fails. Unlike equity investments, the success of a project does not necessarily entitle the funder to any further rights. Unlike research contracts, the funder does not tell the researcher exactly what she wants him to deliver at the end of the research period. Unlike patents, successful grant applications do not confer any right to market exclusivity.

Grant systems also face implementation challenges. Scholars have noted the inefficiency inherent to a system where much of the effort sunk into writing unfunded proposals appears to be wasted (Gross and Bergstrom 2019); they have commented on the unfairness of a system within which female and minority applicants appear to fare less well on average than white, male, or Asian applicants (Ginther et al. 2011); they have provided evidence that peer review sometimes filters out the most novel or creative proposals (Boudreau et al. 2016), or worse, induces scientists to skew their agenda towards projects more likely to generate results in the short term (Azoulay et al. 2011).

Why, then, do grants exist?

We argue that grants are likely to be the most effective—and feasible—way to fund basic research when two fundamental conditions simultaneously hold. First, when the social value of a scientific finding likely exceeds its privately appropriable value. Second, when specifying the parameters of a desired research solution ahead of time is impossible. These twin conditions would appear to characterize much exploratory and early-stage research that is often labeled “basic” or “pure.” We will also discuss two subsidiary arguments in favor grant funding over alternative mechanisms: when potential research performers face financial constraints, and when investments take the form of general-purpose research infrastructure (as opposed to specific projects).

² To be sure, the connection between investments in research and rising living standards or improved national defense has come under increasing political scrutiny (Brooks 1996). Economists and other social scientists have also developed a more nuanced understanding of the innovation process. Over time, they have come to challenge the assumption that for-profit firms would never invest in basic research (Rosenberg 1990), they have incorporated the complex motivations that often guide scientists in their theoretical and empirical studies (e.g., Dasgupta and David 1994; Azoulay et al. 2011), and they have questioned the validity of the distinction between “pure” and applied research in the first place (Stokes 1997). However, this improved understanding does not overturn Arrow and Nelson’s basic insight—that the free market is unlikely to provide the necessary resources for the conduct of scientific research (Balconi et al. 2010).

Limited or undesirable appropriability. There are many cases in which the value that innovations generate for society vastly exceeds what its inventor can be paid. Consider again the case of Gleevec. In addition to being a scientific breakthrough, this molecule was also a financial blockbuster for Novartis, reaching a peak of \$4.65 billion in revenue in 2015 prior to generic entry. Did the promise of such rewards under the patent system provide Novartis with sufficient incentives to develop Gleevec? While Novartis did indeed invest considerable resources in R&D once a candidate drug molecule had been identified, the vast majority of research investments that made Gleevec possible were made long before Novartis started development activities, indeed long before the idea of a treatment approach for CML even existed (Hunter 2007).

These foundational R&D investments included grants made in the 1960s for exploring the genetic basis of cancer, as well as grants made in 1980s for the study of vascular disease. Investment in this type of knowledge is unlikely to be privately profitable: at the time the firm needs to allocate resources for a research project, there is no clear hypothesis for how it would lead to a commercializable drug, meaning that the investment would entail considerable risk for a very small chance of success. Further, even if this research did lead to a testable hypothesis in the context of drug development, the firm making this investment would be enabling other firms to build on this knowledge (for free) to develop their own (competing) drugs.³

Patents, by granting firms a period of market exclusivity, arise as a natural tool to restore innovation incentives. However, an important cost of relying on patent incentives to fund innovation is that they create after-the-fact market distortions through monopoly pricing: when firms have IP protections over their inventions, they will charge a higher price to would-be users relative to a competitive market. In recent years, a growing number of extraordinarily expensive drugs have heightened the salience of this tradeoff: in 2019, for instance, the FDA approved Zolgensma, a gene therapy for a rare childhood disorder, that costs \$2.1 million per patient. While critics argue that such prices are tantamount to extortion, drug makers counter that they are necessary to compensate for the substantial risks of the R&D process. Similarly, although less attention has been paid to more modestly priced drugs, the aggregation of smaller mark-ups on common drugs can also limit access for poorer households and the financial health of ultimate payers such as Medicare.

Open-ended search and contractability. When patents are not appropriate, why not use prizes instead? Research prizes—awarded to whomever achieves a certain outcome first—have several advantages relative to grants, the most obvious of which is the fact that prizes do not need to be paid unless research is successful. In addition, using prizes means that funders do not need to select winners before evaluating their work, making it possible to incentivize research effort from a much larger group of participants (Murray et al. 2012).

³ A similar concern applies when considering innovations targeting the poor, such as treatments for malaria: while there is doubtless social value in addressing the problem (given the massive toll on human health exacted by this disease, particularly in Sub-Saharan Africa and South Asia), neither patients nor their cash-strapped governments can afford to pay for solutions. In light of this reality, firms allocate their R&D resources toward challenges faced by wealthier consumers, who gave both the ability and the willingness to pay for the fruits of innovation.

For example, in 2006, the company Netflix announced an open competition with a \$1 million prize to any team that could improve its recommendation algorithm, the feature that allows to guide users towards movies they are likely to appreciate, thereby boosting willingness to pay for the service. This contest drew entries from over 2,000 teams, a level of participation that would be impossible under a grant model in which winners are selected before research even begins.

However, the structure of the Netflix Prize makes it impossible to replicate in many other research settings. Netflix provided entrants with a large training dataset and was able to articulate a precise, unidimensional metric for assessing both final and intermediate progress (improvement in the root mean-squared error over its current algorithm). This set of parameters were spelled out entirely at the outset, providing contestants with clarity and transparency (Lakhani et al. 2014).

Yet, in many other situations, it would be impossible for a funder to spell out the conditions for winning before seeing any submissions, or to commit to a single or narrow-set of metrics to evaluate success. In the context of exploratory research, narrowing the question in a way that makes it easier to specify, or forcing a solution pathway on potential participants might ultimately stifle innovation and result in sub-optimal solutions.

A related problem with directed search is that the value of particular research results may not appear initially obvious, as in the case of *thermus aquaticus*, the NSF-supported discovery—in Yellowstone National Park—of a bacteria that retains its enzymatic properties under extremely different temperature conditions (Brock and Freeze 1969). Indeed, this is the type of project that might easily have been singled out as an exemplar of wasteful scientific spending—that is, until Kary Mullis and the Cetus Corporation leveraged its unusual properties to develop the polymerase chain reaction (PCR) in the late 1980s, ushering in a new era in biotechnology with applications in far-flung domains such as forensics and paternity testing (Stern 2004).

The constraint on *ex ante* problem formulation suggests that the range of challenges for which innovation contests will dominate other contractual mechanisms, including grants, is perhaps narrower than their proponents have been willing to acknowledge.

Together, appropriability conditions and the nature of idea search are dimensions that can guide policy makers as they navigate the landscape of institutions supporting the production of scientific knowledge.

Figure 1: The Research Funding Ecosystem

		<u>Appropriability</u>	
		Feasible/Desirable	Unfeasible/Undesirable
<u>Nature of Idea Search</u>	Open-ended	Patents	Grants
	Known end-point	Research Contracts	Prizes

As depicted in Figure 1, grants are most suitable in the upper-right quadrant, when appropriating the market returns associated with knowledge production is either infeasible or undesirable, and when the formulation of problems worth solving cannot be scripted in advance. Patents share with grants the ability to harness scientific or technological creativity in a decentralized way, but differ from them in relying on market incentives to stimulate and direct investments. Like grants, prizes promise to direct innovation efforts towards aims that the market might neglect if left to its own devices. Unlike grants, this mechanism requires advanced specifications of the problems worth addressing in order to be effective. Finally, research contracts might operate best in environments where the “deliverable” can be well-specified and appropriability concerns do not loom large (such as in the defense context, with one large paying customer able to specify objectives, with associated penalties for non-performance).

We end this section with two additional arguments that might sometimes push patrons of science to favor grant funding over alternative mechanisms.

Financial constraints. Grants may be particularly effective in cases where researchers are financially constrained. Patents and prizes reward innovators after they have invested in R&D when R&D efforts turn out to be successful. By design, then, this requires innovators to put up capital and bear substantial risk upfront. This is likely to limit both the set of people and organizations who can afford to engage in R&D, and the nature of the R&D they do engage in. While debt and equity markets exist, a large literature in finance suggests that financial frictions nonetheless lead firms, large and small, to underinvest in innovation generally, and in high-risk projects in particular (Froot, Scharfstein, and Stein 1993; Howell 2017, Krieger et al. 2018, Nanda and Rhodes-Kropf 2016). VC investors routinely refer to “financing risk” to describe how otherwise-sound projects may fail to obtain additional capital for continued exploration. The high cost of designing and running experiments that sufficiently reduce uncertainty likely explain why venture capital activity has been circumscribed to a narrow range of sectors (Kerr and Nanda 2015).

While limiting for firms, financial constraints become prohibitive for individual scientists seeking to finance their investigations, especially in fields that necessitate specialized capital equipment (as in condensed matter physics), or expensive materials (such as mice with a particular genetic profile). Without grants, it would be impossible for junior scientists to establish their laboratories and independent research identities. While some universities can afford to provide generous “start-up packages” to their new employees, most institutions are limited in their ability to support researchers absent external grants (Stephan 2012). Relying on localized funding of this type may widen disparities in science, hampering the opportunities available to those at less wealthy institutions.

Supporting human capital and other research “infrastructure.” Because the potential of a given research trajectory is difficult to predict and can shift over time, investments in specific research projects may have less durable payoffs than investments in research “infrastructure,” whether in the form of physical or human capital.

Patents, contracts, and research prizes are not useful tools in this regard because they are necessarily directed to specific ends. Grants, however, are more flexible. While they are frequently used to fund projects (as in the case of the NIH R01), they can also be used to fund institutions (such as when the DOE funds the construction of a new light source at the synchrotron located at Brookhaven National Laboratory), or public goods (e.g., the Sloan Foundation underwriting the Digital Sky Survey, which has created detailed, open access, three-dimensional maps of the universe).

Investments in scientific training and apprenticeship are also typically financed through grants. For instance, Stanford graduate student Sergei Brin was supported by a dissertation fellowship from the National Science Foundation (NSF) when he teamed up with fellow graduate student Larry Page to design *BackRub*, a prototype world wide web search engine that leveraged hyperlinks between pages to develop an “importance” ranking for a set of 24 million web pages (Page et al. 1998). By 1998, Page and Brin had obtained funding that allowed them to move their growing operation away from campus and incorporate Google, Inc. (Hart 2004). In the United States at least, nearly all scientific apprenticeships are funded through grants, whether in the form of individual fellowships as above, in the form of training grants awarded to specific institutions, or indirectly as budget items in traditional project grants.

To summarize, we view scientific grants—such as those used in government-sponsored research—as a particularly effective way of supporting research when outcomes are open-ended and when ensuring the broadest spillovers is viewed as a feature rather than a bug. These traits describe a great deal of “basic” or exploratory research settings—the bedrock of the innovation ecosystem.

2. A short history of the scientific grant

Given their importance, how should science funders organize the grantmaking process? In this section, we consider how scientific endeavors have historically been supported, focusing on the origins of the peer review based systems that have come to dominate modern grant-making.

The earliest precursors of the modern scientific grant were patronage systems widely practiced in Europe, Asia, and the Middle East in both the ancient and early modern periods. For scientists like Galileo, for instance, pursuing knowledge in “experimental philosophy” meant securing the support of a wealthy patron, whose generosity was grounded in a mix of utilitarian and status-seeking motivations. Sustaining the interest of a benefactor often came at the price of skewing one’s investigations towards topics the patron found tasteful or prestigious (Westfall 1985).

As capital requirements increased over time, scientists increasingly began to seek public support. In Europe, financial backing took different forms, from the founding of science departments within long-established universities, to the establishment of free-standing “intramural” research institutes—such as the Physikalisch-Technische Reichsanstalt in Germany (Cahan 1982) or the Pasteur Institute in France (Hage and Mote 2010)—where teaching activities did not take place.

The Royal Society’s experience, 1849-1914. The “government grant” administered by the Royal Society is perhaps the most direct ancestor for modern grant systems. Over the 64 years of its existence, 2,316 grants assisted the investigations of 938 scientists. In 1851, it accounted for about 50% of all the funds appropriated by the British parliament in aid of science, declining to 9% on the eve of the First World War when it was terminated (McLeod, 1971). Although its grants were primarily awarded to members of the Society located in and around London, the Royal Society’s selection process eventually came to function like an early form of peer review. After facing initial accusations of bias, the Society reformed its process, leading to the creation of discipline-specific committees with members elected to four-year terms.

Ultimately, the Victorian-era government grant appears to have withered both because of its trustee’s ambivalence about expanding its scope (for fear that a more ample budget would invite the government to meddle in the Royal Society’s affairs) and because of the growing influence of universities. It would take 40 years, and another world war to create a window of opportunity for reinventing the scientific grant, this time on the other side of the Atlantic.

Rise of the philanthropic foundations. Before the second world war, science funding in the United States was dominated by philanthropic foundations such as the Carnegie, Guggenheim, or Rockefeller foundations. The magnitude of the scientific research outlays of the federal government and large industrial firms, such as DuPont, General Electric, and AT&T may have been more significant, but they were not patrons of science. Rather, they designed and performed the research which they paid for.⁴

The scientific foundations were staffed by professional “managers of science” who cultivated a personal network so as to be informed with respect to the scientists and fields worthy of support, but their financial backing targeted institutions (in particular science departments within universities) rather than individual scientists (Kohler 1976). In the early 1930s, the great depression and its associated financial pressures forced the Rockefeller foundation to suspend its institutional grant program and rely instead on “project grants” that amounted to about \$6,700 per year (about \$125,000 adjusted for inflation) for a typical three-year period (Schneider 2015).⁵ However, the similarities between this scheme and modern government grants are superficial. Grant officers did not rely on peer review, nor did they call for applications in open competitions. Rather they appeared to have exercised considerable discretion in selecting winning projects. Unsurprisingly, these informal practices tended to reinforce prevailing biases and power structures in elite science (Barany 2018, 2019).

Post World-War II transition. The investigator-initiated, renewable, peer-reviewed scientific grant emerged in its modern incarnation shortly after WWII, as officials in the US Public Health Service (PHS) maneuvered to transform a wartime strategy to procure specific research products into a broader grant program.

The window of opportunity was the impending expiration of biomedical research contracts awarded by the Office of Scientific Research and Development (OSRD), the federal agency created to coordinate scientific research for military purposes during World War II. After much bureaucratic infighting (Fox 1987), PHS staff secured the transfer of these contracts to NIH and their transmutation into “contract grants,” a term probably chosen to create ambiguity. Using the authority vested in NIH by the Cancer Act of 1944, PHS staff had laid the foundation of scientific grant making by the middle of 1946. This included the creation of sixteen study sections to review the scientific merits of individual applications, overseen by an academic council nominally in charge of deciding on the final list of recipients (Van Slyke 1946); rules governing grantees’ salaries and pensions, as well as the purchase of equipment; an explicit commitment to protect the freedom of investigators as they performed their investigations; and the choice of an 8% overhead rate over the direct costs of grants so as to minimize “unfairness to less wealthy institutions where establishment of

⁴ For instance, the intramural campus of the National Institutes of Health traces its roots back to a one-room “Laboratory of Hygiene” founded in 1887 as part of the Marine Hospital Service.

⁵ Focusing on individual investigators rather than academic departments was met by resistance on the part of the staff in charge of selecting recipients. Alan Gregg, one of the Rockefeller Foundation’s key officers, explicitly argued against them in a 1937 memorandum, stating that a grant operation was tantamount to setting up a “a huge dispensary of chicken feed.” (Schneider 2015: 280). In what may be a prescient cautionary note, the memo pursued: “the hesitant uncertainty of short-term grants all but insults the intelligence if not the sincerity of the recipient and certainly makes a mockery of long-term planning” (Schneider 2015: 309).

research projects would cause an actual burden to administrative operations" (Fox 1987). Over time, additional policies were implemented to complement solicitations on broad topics with more targeted calls for research on specific areas (Myers 2019).

By the late 1940s, the NIH had become preeminent in medical research as a result of its extramural grant program, expending more than half of all federal funds for medical research. It had strong and growing support in Congress and a powerful constituency in the research community, since a majority of recipients appeared to consistently hail from lower-status institutions not represented on study sections (Munger 1960; Strickland 1989). This apparent success explains in large part why when it finally emerged in 1950, the National Science Foundation also chose investigator-initiated grants awarded to university researchers as its primary contractual mechanism, though peer review appears to have initially played a more minor role in its practices, relative to NIH (Baldwin 2018).

Modern developments. Since the 1950s, scientific grants have spread to many other parts of the US federal government (Departments of Energy, Defense, and Agriculture), to some state governments (such as the California Institute for Regenerative Medicine) and the non-profit sector (March of Dimes, American Cancer Association, The Bill & Melinda Gates Foundation, Howard Hughes Medical Institute, The Chan-Zuckerberg Initiative, etc.) Diffusion outside of the United States has been slower. In 2007, the European Union established the European Research Council (ERC) with an initial annual budget of 7.5 billion euros, an organization which shares many of the practices pioneered by NIH and NSF in the United States (König 2017). Interestingly, grants had not figured prominently for the funding of science within individual member countries up until the ERC founding. This suggests that a certain scale was required to justify the costs of administering a peer-review system capable of processing tens of thousands of applications on a yearly basis.

In advanced economies, "extramural" grant systems (such as those operated by NSF or ERC) coexist with "intramural" institutes (such as the National Laboratories in the US, CNRS in France, the Max Planck Institutes in Germany, or RIKEN in Japan) where the allocation of funds is the outcome of layered administrative processes. In a first step, the overall budget for each institute or laboratory is the result of a political process reflecting national priorities, historical allocations, as well as the clout of laboratory leaders with senior civil servants. In a second step, within each institute, a bureaucratic process disburses funds to particular "laboratories," typically headed by a director overseeing medium-sized teams of scientists, technicians, and postdoctoral fellows. Finally, each director has the decision-making power to allocate her budget across specific projects.

To our knowledge, there is no empirical evidence to date that can speak to the relative merits of "extramural" funding through a decentralized, investigator-initiated process versus "intramural" funding filtered through a decision-making hierarchy. It is possible that hierarchs have better information about the relative quality of projects and initiatives pitched to them by the scientists within their institutions, relative to arm's length peer reviewers without access to "soft information." The other side of the coin, of course, is that poor accountability at the top of the hierarchy makes these same leaders susceptible to influence

activities, since the struggle for resources within each institution is necessarily zero-sum. In the rest of this chapter, we restrict our attention to the design of “extramural” grant systems.

3. A guide to designing grant programs

As illustrated in the introduction, grant funding programs have been instrumental in supporting the development of many important innovations. In this section, we explore issues policymakers face when setting up a new—or reforming an already existing—scientific grant system. In particular, we describe the current state of knowledge and highlight open questions pertaining the following elements of such systems: developing goals and expectations, choosing the scope of what is to be supported, selecting amongst applications, monitoring recipients’ activities, supporting translation and commercialization efforts and, finally, evaluating the grant program’s overall impact. These choices can be consequential because, as persuasively documented in Stephan (2012), the availability and nature of grant funding plays an important role in shaping scientists’ careers and their research incentives.

3.1. Developing goals and expectations

Investing in scientific research requires patience and a tolerance for failure. Imagine a \$1 million project with a minuscule 0.00001% chance of leading to a cure for cancer. In practice, relatively few individual organizations have the risk tolerance to spend \$1 million on a single investment that will fail 99.999% of the time.

Yet imagine there are 200,000 such potential projects, all with a 0.00001% chance of success. If their probabilities of success are independent, then these projects collectively represent a \$200 billion dollar investment with an 87% chance of success. Because of the enormous social value of finding a cure for cancer, almost everyone would agree that this would be a worthwhile portfolio of investments. Yet, risk-averse performers are unlikely to invest in any of the component \$1 million dollar investments as stand-alone projects.

As this simple example illustrates, it is important for funders to think of their investments as part of a broader social portfolio of projects, whether they are supported by government agencies, private sector firms, or non-profit entities (Goodin et al. 2016). Even when the failure rate of individual projects is high, the risk inherent to a diverse portfolio of the same projects may be low enough to make the entire effort worthwhile from a social point of view.

This same portfolio logic can be extended to the design of grant-making organizations and processes: it is important to create grant mechanisms that are diverse in the scientific research areas they support, their time horizons and risk preferences, as well as in the expertise and experiences of those who decide how funds are ultimately allocated.

3.2. Program Scope

Funders must first choose what type of research to support. This decision can be thought of as both a choice of both “horizontal” and “vertical” scope. Horizontally, funders need to choose a research domain or set of domains to support, e.g. a set of disease areas. Vertically, funders must decide where in the research “value-chain” to focus, e.g. on early-stage as opposed to “scaling-up” efforts. From a portfolio perspective, it is important that the ecosystem of grant programs covers as much of this space as possible, for example with some funders focusing on established research domains while others launch new areas of inquiry.

One obvious way of contributing to portfolio approach is for funders to seek out intellectual “white spaces,” i.e., areas of the technological landscape that have not, to date, received much public or private attention. However, a key concern with white spaces is that it is often difficult to determine whether there has been little research in an area because scientific opportunities are scarce or because resources are. Indeed, these tend to be self-reinforcing: areas may not receive funding because there has been little progress to date, but that lack of progress may itself result from a persistent lack of support.⁶

Because of this, establishing a new research area may require a dedicated and sustained effort. In 1958, for example, the division of research grants at NIH created a “study section”—one of the many standing committees charged with evaluating the scientific merits of grant proposals—dedicated to genetics. In addition to recruiting distinguished scientists to serve as members, this new genetics study section took it upon itself to define research standards in this emerging domain, through the organization of symposia that resulted in volumes codifying key methodological aspects of genetics research. In the space of 20 years, the amount of applications increased by an order of magnitude (Crow and Owen 2000).

Today, a similar white-space effort might be needed to explore alternative treatments for Alzheimer’s. A long-standing hypothesis in the Alzheimer’s field holds that a protein fragment called beta-amyloid accumulates in the brain, creating neuron-killing clumps that cause this disorder. For many years, NIH funding for Alzheimer’s has primarily focused this amyloid hypothesis, to the detriment of other research streams focused on oxidative stress, neuroinflammation, and another protein called tau (Begley 2019). As drug candidates based on amyloids have repeatedly failed, the Alzheimer’s research community is increasingly seeing the importance of cultivating a diverse set of treatment hypotheses.

Funders who are not willing or do not have the resources to commit to a sustained effort to address research white spaces can have a stronger impact by funding research in already established research areas. Doing so allows them to take advantage of knowledge spillovers. Indeed, one of the hallmarks of knowledge production is that ideas, once produced, can be freely used by others as inputs in their own research efforts.

⁶ “White space” can also exhibit a geographical dimension. Ganguli (2017) studies a grant program funded by George Soros that provided grants to over 28,000 Soviet scientists shortly after the end of the USSR, in an environment where public support of science had all but evaporated. Not only did these grants more than double publications on the margin, they also significantly induced scientists to remain in the science sector.

When a funder supports research in an already active research area, the scientists they fund can have a larger impact by learning from and contributing to the work of other researchers in the same area.

This approach, however, can also lead to excessive duplication of effort, e.g. “priority races” in which different teams of scientists compete to be the first to publish a discovery, often keeping their work secret in the meantime (Hill and Stein 2019). One way to balance a desire to generate spillovers while avoiding duplication is to conceive white spaces not just in terms of research topics but rather in terms of “vertical” research type. For example, the NIH is clearly the dominant funder for biomedical research, especially for “mature basic research,” that is, projects that have generated enough preliminary evidence to prove their conceptual soundness, but which are not necessarily directed toward an immediate application. Given this, new funders may wish to locate their activities “upstream” of NIH by providing seed funding to de-risk very early stage ideas, allowing scientists to generate the preliminary findings necessary to obtain follow-on NIH project funding. Alternatively, they may consider locating “downstream” to support translational infrastructure that helps science make it out of the lab.

3.3. Developing Research Priorities

Having defined the general scope of a grant program, funders must next choose how to set specific research priorities within their domain. Broadly, funders can elect to be “top down” or “mission oriented”—that is, generating priorities internally and then seeking applications related to these priorities—or “bottom up” or “investigator initiated”—that is, allowing applicants to propose their own projects, so that research priorities are determined after the fact.

Both models are used in practice. The family of federal agencies modeled after DARPA typically operate top down. ARPA-E, for example, identified a gap in energy research on materials for semiconductors and responded by designing a funding program called SWITCHES focused the development of high-voltage (approximately 200–2,000 V), high-current power semiconductor devices and circuits that, upon ultimately reaching scale, could offer affordable breakthrough performance in power electronics, in terms of higher efficiencies, higher switching frequencies (and therefore smaller packages), and higher temperature operation (ARPA-E 2013). In such a program, the funder determines the priority area and then solicits applications on that topic.

In contrast, an agency like NIH largely operates bottom up relying on “investigator initiated” grants. Applicants can submit proposals on any of broad range of topics and methods, which will then be peer-reviewed in one of its 178 chartered study sections (e.g., “Synapses, Cytoskeleton and Trafficking,” “Behavioral Genetics and Epidemiology Study Section,” or “Child Psychopathology and Developmental Disabilities.”). In this model, the agency’s research priorities emerge organically through the application and evaluation process, rather than being specified ahead of time.

The relative merits of a mission-oriented versus investigator-initiated approach are a long-standing object of debate in the science policy community (Mazzucato 2018), one not always informed by compelling

empirical evidence. In our view, the appropriate approach depends on the nature of the research funders intend to support. Returning to our 2×2 classification system from Section 1, top-down programs can be better justified when the funder is confident that it knows and can specify the output that it would like. Meanwhile, bottom-up approaches make sense when a funder wants to support the most promising areas of research but lacks the information to identify those areas on their own.

For example, it is unsurprising that top-down models are common at DARPA, which focuses on R&D for technologies that are relevant for defense. As a branch of the Department of Defense—which is the ultimate buyer for many of these research products—DARPA officials are likely to have a good sense of what the DoD’s needs are, making it easier for them to specify research priorities ahead of time.

In contrast, the NIH is charged with funding research that may eventually lead to improvements in health. The time lags between initial R&D can be long, running into the decades (Li et al. 2017). In this case, it is unlikely that NIH administrators will be informed enough to accurately identify and solicit applications in the most promising research area. Asking them to pick priority areas may therefore lead to an inefficient allocation of funds (Aghion, Dewatripont, and Stein 2008). In such cases, a bottom-up investigator initiated grant process may do a better job of aggregating the collective wisdom of scientists in the relevant communities.

In practice, of course, there are many hybrid models that attempt to capture the advantages of both approaches. The NIH resorts to Request for Applications (RFAs) to focus the energy of the scientific community on areas that are thought to have been neglected or have fallen in between the interstices at the boundary of its peer-review committees (Sampat 2012). This has also proved a flexible way to respond to congressional pressures to fund research on specific diseases (Godefroy 2011).⁷

Meanwhile, reflecting an awareness that top-down priority setting may lead to inefficient allocations, agencies like DARPA and ARPA-E have a rigorous process of “program peer review” in setting research priorities (Azoulay et al. 2019b). At ARPA-E, for instance, proposed programs need to survive a gauntlet of critiques, some coming from existing program directors, and others coming from leaders in the relevant technical community. Using this feedback, the program manager will refine the problem domain, and only then might the agency director approve the program. This type of iterative review with community feedback is therefore a method through which administrators can attempt to overcome their informational disadvantage in identifying high potential research priorities.

In summary, and echoing our earlier point about research portfolios, society might be best served by an ecosystem of funders, some of which set specific agendas in cases when research priorities are clear, and

⁷ One note of caution concerns the potential difficulty of convincing scientists to shift their work into new areas via specific RFPs. Recent research by Myers (2019) suggests that established scientists are relatively “inelastic” in the sense that they are unlikely to switch their research interests in response to small amounts or a small likelihood of funding. Myers’ research suggests, instead, that it would be cheaper to target funds for research in new areas to younger scientists who are more flexible in their research interests.

others that embrace the wide interests of their relevant scientific communities when research goals are more exploratory.

3.4. Grant Evaluation: Peer Review and the Determination of “Scientific Merit”

Once a pool of applications has been collected, grant agencies must select winners and losers. Under the traditional model of peer review, applications are read and scored by multiple evaluators, who then discuss and vote on which applications to fund. This raises three important design questions: who should evaluate the proposals, what types of evaluative input should funders seek from evaluators, and how should these potentially divergent signals be aggregated? We discuss each of these in turn.

Choosing evaluators. First, what kinds of people should grant funders seek advice from when evaluating grant applications? While subject matter experts may have better information about the quality of an application, they may also have preferences—supporting their field, handicapping a competitor—that impede their objectivity. Similarly, reviewers with technical expertise may have a better sense of a project’s feasibility but those with industry or policy expertise may have a better sense of its potential. In a study of NIH peer review, Li (2017) shows that scientists are biased in favor of applicants in their own fields, they are also substantially better informed. Rather than striving to eliminate conflicts of interest entirely, funders should balance potential for bias against the value of an expert’s information.

The determination and use of scientific merit scores. In addition to seeking advice from subject human experts, how much should funders rely on quantitative metrics like publications and citations? Recent empirical studies have shown that following algorithmic advice or other quantitative “rules” may yield better outcomes: Kleinberg et al. (2018) shows that algorithms may be better at predicting recidivism among arrestees; Hoffman, Kahn and Li (2017) show that following algorithmic job test recommendations yields better outcomes than relying on the opinions of human recruiters. This evidence concords with an older stream of research in psychology that compares “clinical” and “actuarial” approaches to decision making and typically finds the latter to be associated with superior outcomes (Dawes et al. 1989).

These studies, however, focus on predicting traits—a worker’s job tenure, for instance—that involve less creativity and variability than assessing scientific potential. Li and Agha (2015) show that human review scores predict eventual research output above and beyond what would be suggested by quantitative metrics alone. In this case, funders should focus on understanding the comparative advantage of human and metrics-based assessments. In the case of the NIH, Li and Agha (2015) show that the relative contribution of humans relative to quantitative metrics is higher amongst top scoring applications.⁸ This suggests a policy

⁸ Cole et al. (1981) and Pier et al. (2018) cast doubt on the fidelity of peer evaluators after finding low rates of agreement between reviewers assessing the merits of the same NSF or NIH grant applications.

in which quantitative metrics can be used to make initial screens, allowing peer reviewers to focus their expertise on distinguishing amongst top performers with a higher chance of being funded.

Finally, funders must also decide how rigidly to adhere to the funding recommendations of external reviewers. Most peer review systems allow for projects to be funded “out of order”—that is, they allow program administrators promote or demote specific projects when doing so would enable the agency to pursue a specific priority.⁹ Ginther and Heggeness (2020) study the careers of applicants to a postdoctoral fellowship program at NIH, and find that “promoted” applicants (those who scored below a cutoff but were nonetheless funded) secure less research funding in the long run, relative to applicants who were “passed over” (i.e., scored above the cutoff but not appointed). It is of course possible that these applicants fared better on other metrics, but at the very least, this evidence should convince agencies to carefully record instances when they choose to deviate from typical funding rules, and track the outcomes that result over time.

Aggregating opinions. Given a chosen set of evaluators (human or otherwise), how should organizations aggregate potentially disparate opinions? The most common approach is to simply take an average; this does a good job of capturing reviewers’ overall assessments, but such averaging could plausibly lead to the selection of more conventional and less risky projects. NIH grant applicants often complain that one bad review is enough to torpedo a proposal, even though the most original projects may be more likely to garner negative reviews because they do not fit neatly within established scientific paradigms. Rather, it is possible that diversity of opinion might itself be a marker of creative potential, in which case funders should look closely at grants with a high *variance* in evaluator scores.

A related approach, similar to that used by the Gates Foundation, is to issue reviewers a limited supply of “gold stars.” This forces reviewers to think carefully about how to allocate their stars across projects (Kolev et al. 2019). One could also issue reviewers a limited number of “rotten tomatoes” which have the capacity to sink a proposal. Both these approaches are used in the private sector by venture capitalists (VCs) considering which startup firms to invest in. Malenko, Nanda, and Rhodes-Kropf (2019) survey VC firms on their aggregation practices show that, for early stage investments, VCs often work on an advocacy model in which a startup can be funded as long as one partner is willing to serve as its champion. This advocacy approach prioritizes a project’s upside potential, which can make sense for investments in early stage firms, when capital commitments are relatively low and there is still a great deal of uncertainty about a firm’s potential. For investments in more mature firms, the authors show that majority voting and consensus models are more common. This practice makes it easier for a single partner to block an investment and thereby focuses on minimizing downside risk. This approach might apply to the scientific funding environment in the case of “big science” projects involving large outlays in specialized physical capital.

⁹ In fact, administrative discretion has been an important feature of peer review systems implemented within scientific journals and funding bodies alike, ever since the Royal Society of London instated the practice in 1831 (Moxham & Fyfe 2018). Baldwin (2018) documents how NSF came to place more emphasis on external referee opinions as a strategy to insulate some of its funding decisions from congressional criticism.

Viewed from a portfolio perspective, it is important for funders to select some projects that represent “safer” bets, and others that are higher-impact but potentially riskier. In doing so, funders should strive to match their selection processes to the goals of the program. For example, the NIH may want to consider adopting an advocacy model (ranking based on maximum scores) in their transformative research program, but may want to continue using average scores in evaluating renewals of existing project grants.

3.5. Post-award program management

A funder’s task need not conclude after it has selected award recipients. Rather, funders must decide the extent to which they want ongoing involvement with funded researchers. The pure prize approach, in which funders reward scientists for past successes, requires little to no post-award management. At the other end of the continuum, grant officers can be involved in the choice of collaborators, the determination of intermediate milestones, with ongoing monitoring and possible early termination of the project. Goldstein and Kearney (2020) use internal data from ARPA-E to document that program staff modify projects frequently, especially project timelines, and that these changes are more sensitive to poor performance than to strong performance. They conjecture that such “active project management,” when combined with high upfront risk tolerance, can be used to enhance the productivity of mission-oriented public research funding.

In addition to explicit directives, funders implicitly shape scientists’ research trajectories through their choice of whether and how to conduct reviews for grant renewal. While some programs are explicitly one shot, grants that hold the promise of renewed funding give funders a lever to continue influencing scientists’ research efforts. The majority of life science labs in the US, for instance, rely on continual renewals of NIH grants (which last 3-5 years per cycle) in order to operate. This type of staged funding enables funders to deepen their financial commitment only after ideas have shown some promise. Indeed, staged funding is also standard practice in venture finance: by investing smaller initial amounts, firms can afford to take risks on early stage projects while preserving the option to abandon projects that show no initial promise (see Kerr, Nanda, and Rhodes-Kropf (2014) for an overview of private sector VC financing).

Under such models, scientists have a strong incentive to demonstrate productivity and success in order to renew their funding. These incentives work best when the funder has a clear sense of what behaviors it would like scientists to adopt, a way to measure these outcomes, and is cognizant of the potential for unintended consequences. A renewal policy that emphasizes publication counts, for instance, may lead scientists to waste time on weak projects (or engage in data-mining) in order to seek a publication, rather than accepting initial failures and moving on. Fearing failure, scientists may also take fewer risks initially, steering their work toward safer but potentially less impactful projects.

To address these concerns, organizations that seek to encourage scientific risk-taking must match their rhetoric with deeds. For example, HHMI medical investigators receive funding for an initial period of five years, but the first renewal decision appears rather lax, focusing mostly on whether the funded scientists have made use of the freedom an HHMI investigator-ship allows to branch out in new directions. Azoulay

et al. (2011) show that these failure-tolerant policies influence how these scientists lead their laboratories, the type of personnel they employ, and the methods and questions they choose to investigate. Compared to a matched class of NIH funding recipients (who face a more traditional output-based renewal process), HHMI investigators produce very highly cited publications at a higher rate, as well as more “duds” with few or no citations, which is what one would expect if they chose to privilege “exploration” at the expense of “exploitation” of traditional scientific approaches.

Again, it is important to match post-award management behavior with the program’s goals. For programs that are designed to foster high risk exploratory research, it is important that funders exhibit some flexibility and risk tolerance, especially earlier on. For programs that target progress on more routine research, then it may be more appropriate to adopt renewal policies contingent on reaching defined milestones.

3.6. Translation and Impact

Though grants are mechanisms that enable funders to support basic research, one of their fundamental rationale is that investment in basic science underwrites technological progress through commercialization and other translation efforts (Bush 1945). Yet the majority of academic research supported by agencies like the NSF and NIH do not yield follow-on economic activity in a direct way, whether in the form of patenting, licensing, or entrepreneurship. And, for the subset of ideas that are commercialized, few make it past the so-called “valley of death” to reach a wider audience (Contopoulos-Ioannidis et al. 2003; Beard et al. 2009).

One potential barrier to greater translation is the fact that scientists, left to their own devices, do not necessarily consider engagement with industry as an integral part of their job description (Barham et al. 2020; Cohen et al. 2019). As such, policymakers need to consider the desirability and feasibility of incorporating “translation incentives” in the design of grant systems. As an approximation, it is useful to distinguish between a passive approach—whereby obstacles to commercialization (such as unclear or limited IP rights) are removed—from an active approach—whereby funders are directly involved in helping their awardees commercialize their research.

Passive translation: IP rights and grant policy. In the United States, the Bayh-Dole Act (passed in 1980) allows researchers and universities to retain IP rights to inventions supported by federal funding, whereas previously these rights would have in most cases resided with the government. This change contributed to an already growing trend in university patenting and licensing, as documented by Mowery et al. 2001). Part of this increase is credited to organizational investments that universities made in establishing technology transfer offices to facilitate the patenting and licensing of inventions that emerged from academic labs. Implicit in this logic is that academic scientists may lack the knowledge, time, or interest in managing the commercialization of their inventions; they may not know which companies to approach or how to negotiate licensing agreements. Technology transfer offices therefore provide a set of services that complement the scientists’ technical expertise. Reflecting this, universities and scientists typically split

revenues associated with an invention, although the extent to which academics respond to the level of this royalty rate is in dispute (Hvide and Jones 2018; Ouelette and Tutt 2020).

Hausman (2019) studies the impact of Bayh-Dole on measures of real economic activity, in order to better understand the role that university science plays in shaping invention and entrepreneurship in the local economy. She finds that employment, wages, and corporate innovation appear to increase as a result of Bayh-Dole: these measures of economic output rise more rapidly after Bayh Dole in counties near universities and in industries more closely related to the local university's areas of innovative expertise.

However, a key critique of Bayh-Dole (and other IP-rights focused policies) is that an increased focus on patenting may weaken universities' commitments to "open science." Williams (2013) and Murray et. al. (2016) both consider the value of open access in scientific research. Williams focuses on IP rights related to human genes and finds that genes sequenced by the private firm Celera, and therefore subject to its IP, were less likely to be the subject of follow on research and product development, relative to comparable open-access genes sequenced by the Human Genome Project. Murray et al. (2016) further examine how IP rights shape the nature of the follow on research that investigators pursue. The authors show that open access to scientific inputs—in this case, genetically engineered mice—encouraged entry by new researchers and lead to a greater diversity of research paths. Together, these and other studies document an IP policy trade-off when policymakers decide whether to allow scientists (and their employers) to patent findings that emerge out of public or even non-profit funding: while strong IP rights provide incentives for the development and commercialization of technologies that would otherwise remain in an embryonic state, these may also limit access for innovators building on the initial work, thereby limiting the scope for non-directed spillovers (Scotchmer 1991; Walsh et al. 2005).

One potential hybrid approach is to allow universities to patent and license their inventions to private sector firms, but to maintain free access to other academic or non-profit users. This would retain the incentive benefits of IP rights but maintaining some commitment to open science.

Active translation and the "ARPA Model." In addition to removing IP barriers, grant funders can take a more active approach toward midwifing the translation of scientific results into prototypes or technologies, as one particular aspect of post-award management mentioned above. This orientation toward commercial impact has been a hallmark of DARPA-style funding, but these efforts have probably been made easier insofar as the Department of Defense is both the funder and ultimate buyer of the inventions that arise from its support. A fairer test of active translation efforts might therefore involve a funder in a domain where technological inputs must be purchased on the open market (Azoulay et al. 2019b).

ARPA-E "tech-to-market" (hereafter "T2M") program and personnel provides a proof-of-concept for active funder involvement, although one that must still be regarded as an ongoing experiment rather than accepted best practice. Before receiving award funds, ARPA-E performers are required to develop a T2M plan in close coordination with ARPA-E's T2M advisors. Commercialization strategies developed to meet this requirement include training and the development of the business information necessary to understand market needs and tailor technology development to address those needs. ARPA-E also helps awardees

develop relationships with relevant government agencies, technology transfer offices, companies, investors, and other organizations to facilitate transition to the commercial phase (National Academies of Sciences, Engineering, and Medicine 2017).

Regardless of the approach espoused by grant system designers, one uncontroversial theme emerges from scholarship on this topic: funders should attempt to lower the cost faced by their awardees while sharing the output of their work with a diverse audience, including other researchers who may produce follow-on work, as well as researchers in industry who may have the expertise and financial wherewithal to develop early-stage ideas and bring them closer to market. One way to do so is for funding agencies to assist in building institutions that make it easier to access materials and knowledge. For instance, in the life sciences, the ability to build on prior research often depends on access to biological specimens—cell lines, tissue cultures, etc. Furman and Stern (2011) demonstrate that biological resource centers, which certify the fidelity of biological materials and facilitate their distribution, substantially amplify the impact of published research, sometimes doubling the number of citations it receives. From the grant funder’s perspective, these types of investments can vastly increase the overall returns to its R&D investments.

4. Toward a science of science funding

Finally, as with any other investment, funders of scientific research should understand the impact that their resources are having. This provides an opportunity to build on strengths in their existing funding model and to improve on weaknesses.

Yet evaluation is difficult without some initial planning. Imagine that a foundation awards a grant to a scientist and two years later she has trained three graduate students, and published 10 additional articles, several of them in prominent journals. In order to assess the impact of this grant, it is not enough to tabulate these outputs, however impressive they can appear. Rather, one needs to understand what her research outcomes would have been like had she not received any support. This is analogous to the challenge that scientists face when assessing the impact of a medical treatment: how does one know whether the patient got better because of the treatment or because of something else?

In medicine, scientists address this challenge by comparing outcomes for treated patients with outcomes for a control group of similar patients who were not treated. Funders of scientific research can do the same by collecting data on similar scientists who were not funded. To begin assessing the value of a grant, one should compare research outcomes between funded and unfunded groups. This comparison is valid if funded and unfunded applicants are similar. If applicants are rejected because they are substantially less qualified, then they would likely have worse research outcomes than funded applicants, even in the absence of funding. Such a comparison would tend to overstate the role of the grant.

The most effective way to address this is to randomize who gets funding. This is akin to randomization in medical trials, or A/B testing in business settings. When applied in the science funding setting, randomized

evaluations seek to determine the impact of grant funding or grant programs by comparing the outcomes of a group that receives funding or is subject to a particular set of grant policies (the treatment group) against outcomes among the group that is not (the control group). Because the program is randomly assigned, members of treatment and control groups do not differ systematically at the start of the evaluation: this allows researchers to attribute any differences in outcomes that may emerge to the causal impact of the grant or grant policy.

Randomized control trials (RCTs) have become the gold standard for policy evaluation and evidence based decision-making. Many governments and foundations use RCTs to assess the efficacy of their programs, and a variety of organizations have emerged, inspired by organizations such as the Poverty Action Lab, to facilitate these experiments. To design an effective and fair RCT evaluation, it is important to appreciate the institutional context and goals at hand. For example, HHMI grants are aimed at encouraging scientists to pursue risky avenues of research, even if this means that in many cases experiments will fail and scientists may have little to publish. In this case, an RCT that focuses on counting publications would be inappropriate because publication counts do not reflect the underlying goal of the organization. For reasons such as these, we believe that the most effective evaluations arise from collaborations between agency staff and external program evaluators.

In many cases, there is reluctance to implement RCTs because of their perceived costs or inefficiencies. Funders, for instance, may understandably not want to randomly allocate their scarce funds to unqualified scientists. Yet, even when a full scale RCT is infeasible, it is still possible to perform some kind of randomization. For example, funders could devise a two-step approach where applicants are first screened to eliminate those that are below a baseline level of acceptable quality; funding could then be randomized within the set of remaining applicants.¹⁰ This would ensure a level of quality control while still enabling funders to better understand the impact of their program.

In addition, there are often other naturally occurring “experiments” that allow researchers to assess the impact of funding. For example, funding cutoffs—so called “pay lines”—create opportunities to use a “regression discontinuity design” where one compares outcomes for those just above and just below the cutoff. The idea is that because their scores are actually quite close, these applicants are likely to be more similar to each other than the average funded applicant is to the average unfunded applicant. As such, differences in their outcomes can be attributed more readily to the grant. Jacob and Lefgren (2011), Howell (2017), and Azoulay et al. (2019a) are all examples of this type of analysis applied to grant funding.

When a fully randomized or “natural” experiment is not possible, an alternative approach is to collect basic data on the characteristics of applicants—for instance, highest education, year of graduation, undergraduate and graduate institution, prior funding history, and keywords describing primary fields of research—and

¹⁰ Fang et al. (2016) propose a modified lottery scheme in this exact spirit.

use these variables to make sure that one is comparing funded and unfunded scientists who look similar in terms of education, past research productivity, and other observable traits.

It is also important to consider the unit of analysis. An individual-level analysis typically yields an estimate of the average effect of being “treated” by funding, that is, the impact of funding for a typical scientist. Funders, however, may be interested in understanding their impact on a field of research as a whole. In this view, it is not enough to compare treatment and control outcomes at the level of an individual scientist because two applicants may have similar ideas: if funding enables one scientist to publish her results ahead of another, that yields a big impact from the perspective of her individual output, but it may not yield as large an impact on her field because that research idea would have been performed regardless. In order to assess the impact of funding on an entire area, one can still apply the same techniques as the ones described above, but focusing on fields rather than individuals as the unit of “treatment.” For example, if one decides to focus funding on translational research in diabetes, one may compare the number of new clinical trials in diabetes to those in other similar disease areas.

Finally, an informative program evaluation requires that funders collect information on research outcomes. While the overall desired impact of a program may be to improve life expectancy for patients with a particular health condition, the long lags involved, as well as the traceability challenges mentioned in the introduction may make it infeasible to deploy metrics that are directly welfare-relevant. In contrast, it may be easier to measure narrower, or intermediate outputs in the innovation process. Before discussing the merits of such “surrogate markers” for impact, it is worth remembering that the outcomes funders track invariably morph into the incentives scientists face. Programs that only track publications (perhaps in “high impact” journals) will provide recipients with an incentive to publish, but may not necessarily stir their interest in seeing their work translated or commercialized. Conversely, funders who carefully tabulate their awardees’ patents may unwittingly lead them to patent unimportant work, as seems to have been the case with patent promotion policies in China (Long and Wang 2018). The most common metrics used in funding program evaluations include:

- Bibliometric measures, such as publications, publications in top journals, or “blockbuster publications,” that is those that receive citations above some absolute threshold (e.g., in the top 1% given their vintage). While not a panacea, these metrics are correlated with subsequent breakthrough discoveries (Lawani and Bayer 1983). These should be considered a basic part of any impact evaluation, even if they can appear far removed from the impact that funders wish to make in their respective domains.¹¹
- Commercial or applied impact. A weakness of publication-based measures is that they may fail to capture the impact of a scientist or research program outside of academia. For mission-oriented organizations in particular, one may want to consider other metrics such as patents

¹¹ A related point is that work on the development and validation of citation-based metric has been a vibrant area of inquiry in the emerging ‘science of science’ field. Recent efforts include attempts to distinguish “consolidating” from “disruptive” publications in science using a combination of backward references and forward acknowledgements (Funk and Owen-Smith 2017; Wu et al. 2019).

- generated (Goldstein and Kearney 2018), clinical trials initiated (Kolev et al. 2019), or the incorporation of start-up firms with growth ambitions (Kearney 2020).
- Career outcomes. Funders may be interested in supporting scientific training rather than specific projects, in which case impact assessments should include measures of career traction or influence: job appointments and promotions, as well as the number and placement of students the researcher trains (Azoulay et al. 2020).

Jaffe (1998) provides a seven-point “wish list” for innovation metrics that science policymakers should have in mind when evaluating the impact of funding programs. First, metrics should have a high signal/noise ratio; second, error in measurements should be uncorrelated with other phenomena of interest; third, the relationship between the proxy and the underlying phenomenon of interest should be linear, or at least of known functional form; fourth, the relationship between the proxy and the underlying concept should be stable over time; fifth, there should be stability across settings (institutional, geographic) in the relationship between the proxy and the underlying concept; sixth, the metric should not be susceptible to easy manipulation or inflation; seventh, it should be possible to consistently track the metric at different levels of aggregation (geographic or institutional).

This list makes for sobering reading since it can be argued that most, if not all of the metrics used in program evaluations to date fall short in at least one respect. This suggests that funders should consider collecting information for a battery of outcomes rather than a single proxy. We also note that the scientific enterprise tends to generate digital breadcrumbs that, when systematically collected and parsed, can help alleviate traceability challenges and narrow the gap between bibliometric data and welfare-relevant outcomes. For instance, the wide availability of genetic sequence information as metadata attached to publications has made it possible for researchers to trace the impact of basic genetics research from the laboratory all the way to clinical trials, and the market availability of diagnostic tests (Williams 2013; Kao 2020).

In addition to impact evaluation, there could be significant returns to examining design elements of the funding system. Is scientific funding more effective when it holds scientists accountable for the precise content of the investigations they proposed (as is the case for NIH and NSF), or when it gives them the flexibility to alter the content of their research in the middle of a funding cycle (as is the case in the HHMI investigator program)? Should evaluator sentiment be averaged to generate priority scores, or can quadratic voting approaches be used to incorporate the intensity and variance in evaluator sentiment when scoring proposals? Should young and established investigators’ proposals be evaluated in the same pool, or on separate tracks? These are empirical questions whose answers can only be provided through the careful design of tailored experiments.

Given the high potential returns to evaluation and experimentation, we end this section by pondering why the scientific community, funding agencies, and non-profit foundations have been so reluctant to “turn the scientific method on themselves” (Azoulay 2012). Conservatism on the part of those benefiting from the status quo certainly plays a role, but resistance to experimentation does not only reflect self-serving motives. First, there are objective obstacles to experimentation in this setting, namely, the long lags involved for welfare-relevant outcomes to be realized, and the scale required to power experiments in order to detect

meaningful differences in a world where “tail” outcomes are inherently more informative than “average” outcomes.

Second, science policy makers might fear the nuanced implications from careful analysis might open the door to budgetary restrictions, whereas the emphasis on carefully cherry-picked anecdotes does not entail a similar degree of political risk. Paradoxically, the routinization of experimentation in scientific funding might require the imposition of a mandate from political institutions.

5. Conclusion

The investigator-initiated scientific grant is an important meta-institution with distinctly American origins, and one of the touchstones of the US “National Innovation System” (Nelson 1993). Yet, it would be surprising if the initial design choices made by institutional entrepreneurs such as Vannevar Bush and C. James van Slyke in 1945 continued to provide a comprehensive blueprint for policy makers seeking to meet the challenges of scientific discovery in the 21st century.

While this essay has attempted to grapple with some of the delicate trade-offs present in the design of science funding institutions, we end the chapter by emphasizing a small number of core principles for policymakers.

First, there is great worth in maintaining a diversity of approaches for grant making. The analysis of grant systems should therefore be approached as a portfolio evaluation problem. A crucial activity for science policymakers is therefore the identification of gaps in the ecosystem of funding. Traditionally, topic “white space” have been most salient, but we believe that it could be at least as productive to identify gaps with respect to risk-orientation. As an example, while it would not be prudent to encourage federal agencies to adopt HHMI-like funding practices wholesale, at present neither NIH nor NSF have in their arsenal a mechanism providing grantees with a truly long-term horizon to plan their investigations (e.g., seven to ten years).¹²

Second, in the non-profit and public sector alike, funders have proved surprisingly reluctant to submit changes in the administration of their grants to rigorous evaluation. Nor do funders typically routinize the collection of outcome information regarding the applicants they did not choose to support. The lack of an experimental mindset partly explains why so many important questions regarding the design of grant systems remain without clear answers, and also why specific advice provided to policymakers must be tempered. Rather than chase the latest funding fad (“people, not projects,” a modified funding lottery, a “translational” institute, replacing grants with prizes, etc.), turning the scientific method on the funding process could yield novel insights with the potential to accelerate scientific discoveries (Azoulay 2012).

¹² Recently, the National Institute of General Medical Science (NIGMS), NIH’s component institute focused on “basic” biological research initiated the R35 “Maximizing Investigators’ Research Award” (MIRA) which is a step in this direction, though the time horizon of these awards is only five years.

Within this framework, federal funding agencies and philanthropic funders could encourage randomized experimentation of grant-making practices—whether they pertain to peer review, time horizon, or intellectual property policies—and carefully evaluate the results before adopting them at scale.

In sum, scientific grant funding is an important part of the policy toolkit for encouraging innovation, particularly in basic research. In this chapter, we have covered a range of examples—from the NIH to the NSF to the DOD and DOE—of agencies that have used varying types of grant mechanisms to support both incremental and high-risk R&D. By adopting a more scientific approach to studying the grant funding processes, policymakers can refine these tools to support new research challenges and needs.

References

Aghion, Philippe, Mathias Dewatripont, and Jeremy C. Stein. 2008. "Academic Freedom, Private Sector Focus, and the Process of Innovation." *RAND Journal of Economics* 39(3): 617-635.

ARPA-E. 2013. "Final Assistance Funding Opportunity Announcement, Strategies for Wide-Bandgap, Inexpensive Transistors for Controlling High Efficiency Systems (SWITCHES). DE-FOA-0000942." Department of Energy, Washington, DC.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Invention." In, *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pp. 609-625. Princeton, NJ: Princeton University Press.

Azoulay, Pierre. 2012. "Turn the Scientific Method on Ourselves." *Nature* 484(7392): 31-32.

Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso. 2011. "Incentives and Creativity: Evidence from the Academic Life Sciences." *RAND Journal of Economics* 42(3): 527-554.

Azoulay, Pierre, Joshua S. Graff-Zivin, Danielle Li, and Bhaven N. Sampat. 2019a. "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules." *The Review of Economic Studies* 86(1): 117-152.

Azoulay, Pierre, Erica Fuchs, Anna P. Goldstein, and Michael Kearney. 2019b. "Funding Breakthrough Research: Promises and Challenges of the 'ARPA Model'." *Innovation Policy and the Economy* 19: 69-96.

Azoulay, Pierre, Wesley H. Greenblatt, and Misty L. Heggeness. 2020. "Long-Term Effects from Early Exposure to Research: Evidence from the NIH 'Yellow Berets'." NBER Working Paper #26069.

Balconi, Margherita, Stefano Brusoni, and Luigi Orsenigo. 2010. "In Defence of the Linear Model: An Essay." *Research Policy* 39(1): 1-13.

Baldwin, Melinda. 2018. "Scientific Autonomy, Public Accountability, and the Rise of "Peer Review" in the Cold War United States." *Isis* 109(3): 538-558.

Barany, Michael J. 2018. "A Postwar Guide to Winning a Science Grant." *Physics Today*, 20 March 2018.

Barany, Michael J. 2019. "Rockefeller Bureaucracy and Circumknowing Science in the Mid-twentieth Century." *International Journal for History, Culture and Modernity* 7: 779-796.

Barham, Bradford L., Jeremy D. Foltz, and Ana Paula Melo. 2020. "Academic Engagement, Commercialization, and Scholarship: Empirical Evidence from Agricultural and Life Scientists at U.S. Land-Grant Universities." NBER Working Paper #26688.

Beard, T. Randolph, George S. Ford, Thomas M. Koutsky, and Lawrence J. Spiwak. 2009. "A Valley of Death in the Innovation Sequence: An Economic Investigation." *Research Evaluation* 18(5): 343-356.

Begley, Sharon. 2019. "The Maddening Saga of How an Alzheimer's 'Cabal' Thwarted Progress Toward a Cure for Decades." *STAT*, June 25th, 2019.

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, and 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.

Brock, Thomas D., and Hudson Freeze. 1969. "Thermus aquaticus gen. n. and sp. n., a Nonsporulating Extreme Thermophile." *Journal of Bacteriology* 98(1): 289-297.

Brooks, Harvey. 1996. "The Evolution of U.S. Science Policy." In Bruce L.R. Smith, and Claude E. Barfield (Eds.), *Technology, R&D, and the Economy*, pp. 15-48. Washington, DC: The Brookings Institution.

Bush, Vannevar. 1945. *Science: The Endless Frontier*. Washington, DC: US General Printing Office.

Cahan, David. 1982. "Werner Siemens and the Origin of the Physikalisch-Technische Reichsanstalt, 1872-1887." *Historical Studies in the Physical Sciences* 12(2): 253-283.

Cole, Stephen, Jonathan R. Cole, and Gary A. Simon. 1981. "Chance and Consensus in Peer Review." *Science* 214(4523): 559-567.

Contopoulos-Ioannidis, Despina G., Evangelia E. Ntzani, and John P.A. Ioannidis. 2003. "Translation of Highly Promising Basic Science Research into Clinical Applications." *The American Journal of Medicine* 114(6): 477-484.

Cohen, Wesley M., Henry Sauermann, and Paula Stephan. 2019. "Not in the Job Description: The Commercial Activities of Academic Scientists and Engineers." NBER Working Paper #24769.

Crow, James F., and Ray D. Owen. 2000. "Kay Wilson and the NIH Genetics Study Section." *Genetics* 155(1): 1-5.

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

Dawes, Robyn, David Faust, and Paul E. Meehl. 1989. "Clinical Versus Actuarial Judgment." *Science* 243(4899): 1668-1674.

Fang, Ferric C., and Arturo Casadevall. 2016. "Research Funding: the Case for a Modified Lottery." *mBio* 7(2): e00422-16.

Fox, Daniel M. 1987. "The Politics of the NIH Extramural Program, 1937-1950." *Journal of the History of Medicine & Allied Sciences* 42(4): 447-466.

Froot, Kenneth A., David S. Scharfstein , and Jeremy C. Stein. 1993. "Risk Management: Coordinating Corporate Investment and Financing Policies." *The Journal of Finance* 48(5): 1629-1658.

Funk, Russell J., and Jason Owen-Smith. 2017. "A Dynamic Network Measure of Technological Change." *Management Science* 63(3): 791-817.

Furman, Jeffrey, and Scott Stern. 2011. "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Knowledge Production." *American Economic Review* **101**(5): 1933-1963.

Gallini, Nancy, and Suzanne Scotchmer. 2002. "Intellectual Property: What is the Best Incentive System?" *Innovation Policy and the Economy* **2**: 51-77.

Ganguli, Ina. 2017. "Saving Soviet Science: The Impact of Grants When Government R&D Funding Disappears." *American Economic Journal: Applied Economics* **9**(2): 165-201.

Ginther, Donna K., Walter T. Schaffer, Joshua Schnell, Beth Masimore, Faye Liu, Laurel L. Haak, and Raynard Kington. 2011. "Race, Ethnicity, and NIH Research Awards." *Science* **333**(6045): 1015-1019.

Ginther, Donna K., and Misty L. Heggeness. 2020. "Administrative Discretion in Scientific Funding: Evidence from a Prestigious Postdoctoral Training Program." *Research Policy* **49**(4): 103953.

Godefroy, Raphael. 2011. "The Birth of the Congressional Clinic." Working Paper, Paris School of Economics.

Goldstein, Anna P., and Michael Kearney. 2018. "Uncertainty and Individual Discretion in Allocating Research Funds." Available at SSRN: <https://ssrn.com/abstract=3012169>.

Goldstein, Anna P., and Michael Kearney. 2020. "Know When to Fold 'em: An Empirical Description of Risk Management in Public Research Funding." *Research Policy* **49**(1): 103873.

Goodin, Michael M., Graham F. Hatfull, and Harmit S. Malik. 2016. "A Diversified Portfolio." *Annual Review of Virology* **3**: vi-viii.

Gross, Kevin, and Carl T. Bergstrom. 2019. "Contest Models Highlight Inherent Inefficiencies of Scientific Funding Competitions." *PLoS Biology* **17**(1): e3000065.

Hage, Jerald, and Jonathon Mote. 2010. "Transformational Organizations and a Burst of Scientific Breakthroughs: The Institut Pasteur and Biomedicine, 1889-1919." *Social Science History* **34**(1): 13-46.

Hart, David. 2004. "On the Origins of Google." National Science Foundation. Accessed on February 17th, 2020 at https://www.nsf.gov/discoveries/disc_summ.jsp?cntn_id=100660.

Hausman, Naomi. 2019. "University Innovation and Local Economic Growth." Working Paper, Hebrew University of Jerusalem.

Hill, Ryan and Carolyn Stein. 2019. "Scooped! Estimating Rewards for Priority in Science" Working Paper, Massachusetts Institute of Technology.

Hoffman, Mitchell, Lisa B. Kahn, and Danielle Li. 2018. "Discretion in Hiring." *The Quarterly Journal of Economics* **133**(2): 765-800.

Howell, Sabrina T. 2017. "Financing Innovation: Evidence from R&D Grants." *American Economic Review* **107**(4): 1136-1164.

Hunter, Tony. 2007. "Treatment for Chronic Myelogenous Leukemia: The Long Road to imatinib." *Journal of Clinical Investigation* **117**(8): 2036-2043.

Hvide, Hans K., and Benjamin F. Jones. 2018. "University Innovation and the Professor's Privilege." *American Economic Review* **108**(7): 1860-1898.

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

Jaffe, Adam. 1998. "Measurement Issues." In Lewis Branscomb, and James Keller (Eds.), *Investing in Innovation: Creating a Research and Innovation Policy That Works*, pp. 64-84. Cambridge, MA: The MIT Press.

Kao, Jennifer. 2020. "Charted Territory: Evidence from Mapping the Cancer Genome and R&D Decisions in the Pharmaceutical Industry." Working Paper, UCLA.

Kearney, Michael. 2020. "Translating Science through Startups: Evidence from the National Science Foundation's I-Corps Program." Working Paper, Massachusetts Institute of Technology.

Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as Experimentation." *Journal of Economic Perspectives* 28(3): 25-48.

Kerr, William R., and Ramana Nanda. 2015. "Financing Innovation." *Annual Review of Financial Economics* 7: 445-462.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. "Human Decisions and Machine Predictions." *The Quarterly Journal of Economics* 133(1): 237-293.

Kohler, Robert E. 1976. "The Management of Science: The Experience of Warren Weaver and the Rockefeller Foundation Programme in Molecular Biology." *Minerva* 14(3): 279-306.

Kolev, Julian, Pierre Azoulay, Yuly Fuentes-Medel, and Fiona Murray. 2019. "Expert Evaluation in Innovation: The Role of Distance and Consensus in Project Selection." Working Paper, Southern Methodist University.

König, Thomas. 2017. *The European Research Council*. Cambridge, UK: Polity Press.

Krieger, Joshua, Danielle Li, and Dimitris Papanikolaou. 2018. "Missing Novelty in Drug Development." NBER Working Paper #24595.

Laffont, Jean-Jacques, and Jean Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: The MIT Press.

Lakhani, Karim R., Wesley M. Cohen, Kynon Ingram, Tushar Kothalkar, Maxim Kuzemchenko, Santosh Malik, Cynthia Meyn, Greta Friar, and Stephanie Healy Pokrywa. 2014. "Netflix: Designing the Netflix Prize (A)." Harvard Business School Case 615-015.

Lawani, Stephen M., and Alan E. Bayer. 1983. "Validity of Citation Criteria for Assessing the Influence of Scientific Publications: New Evidence with Peer Assessment." *Journal of the American Society for Information Science* 34(1): 59-66.

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

Li, Danielle, Pierre Azoulay, and Bhaven N. Sampat. 2017. "The Applied Value of Public Investments in Biomedical Research." *Science* 356(6333): 78-81.

Li, Danielle, and Leila Agha. 2015. "Big Names or Big Ideas: Do Peer-review Panels Select the Best Science Proposals?" *Science* 348(6233): 434-438.

Long, Cheryl Xiaoning, and Jun Wang. 2019. "China's Patent Promotion Policies and its Quality Implications." *Science and Public Policy* 46(1): 91-104.

MacLeod, R. M. 1971. "The Royal Society and the Government Grant: Notes on the Administration of Scientific Research, 1849-1914." *The Historical Journal* 14(2): 323-358.

Malenko, Andrey, Ramana Nanda, and Matthew Rhodes-Kropf. 2019. "Investment Committee Voting and the Financing of Innovation." Working Paper, Boston College.

Mazzucato, Mariana. 2018. "Mission-oriented Innovation Policies: Challenges and Opportunities." *Industrial and Corporate Change* 27(5): 803-815.

Mokyr, Joel. 2002. *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton, NJ: Princeton University Press.

Mowery, David C., Richard R. Nelson, Bhaven N. Sampat, and Arvids Ziedonis. 2001. "The Growth of Patenting and Licensing by U.S. Universities: An Assessment of the Effects of the Bayh-Dole Act of 1980." *Research Policy* 30(1): 99-119.

Moxham, Noah, and Aileen Fyfe. 2018. "The Royal Society and the Prehistory of Peer Review." *The Historical Journal* 61(4): 863-889.

Munger, Mary G. 1960. Growth of the External Programs of the National Institutes of Health. Statistics and Analysis Branch, Division of Research Grants, National Institutes of Health.

Murray, Fiona, Philippe Aghion, Mathias Dewatripont, Julian Kolev, and Scott Stern. 2016. "Of Mice and Academics: Examining the Effect of Openness on Innovation." *American Economic Journal: Economic Policy* 8(1): 212-252.

Murray, Fiona, Scott Stern, Georgina Campbell, and Alan MacCormack. 2012. "Grand Innovation Prizes: A Theoretical, Normative, and Empirical Evaluation." *Research Policy* 41(10): 1779-1792.

Myers, Kyle. 2020. "The Elasticity of Science." Forthcoming, *American Economic Journal: Applied Economics*.

Nanda, Ramana, and Matthew Rhodes-Kropf. 2016. "Financing Entrepreneurial Experimentation." *Innovation Policy and the Economy* 16: 1-23.

National Academies of Science, Engineering, and Medicine. 2017. *An Assessment of ARPA-E*. Washington, DC: The National Academies Press.

Nelson, Richard R. 1959. "The Simple Economics of Basic Scientific Research." *Journal of Political Economy* 67(2): 297-306.

Nelson, Richard R. 1993. *National Innovation Systems*. New York: Oxford University Press.

Ouellette, Lisa Larrimore, and Andrew Tutt. 2020. "How Do Patent Incentives Affect University Researchers?" *International Review of Law and Economics*.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1998. "The PageRank Citation Ranking: Bringing Order to the Web." Working paper, Stanford University.

Pier, Elizabeth L., Markus Brauer, Amarette Filut, Anna Kaatz, Joshua Raclaw, Mitchell J. Nathan, Cecilia E. Ford, and Molly Carnes. 2018. "Low Agreement Among Reviewers Evaluating the Same NIH Grant Applications." *Proceedings of the National Academy of Sciences* 115(12): 2952-2957.

Price II, W. Nicholson. 2018. "Grants." *Berkeley Technology Law Journal* 34(1): 1-66.

Rivest, Ronald L., Adi Shamir, and Leonard Adleman. 1978. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems." *Communications of the ACM* **21**(2): 120-126.

Rosenberg, Nathan. 1979. "Technological Interdependence in the American Economy." *Technology and Culture* **20**(1): 25-50.

Rosenberg, Nathan. 1990. "Why Do Firms Do Basic Research (With Their Own Money)?" *Research Policy* **19**(2): 165-174.

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

Schneider, William H. 2015. "The Origin of the Medical Research Grant in the United States: The Rockefeller Foundation and the NIH Extramural Funding Program." *Journal of the History of Medicine & Allied Sciences* **70**(2): 279-311.

Scotchmer, Suzanne. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives* **5**(1): 29-41.

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

Stern, Scott. 2004. *Biological Resource Centers: Knowledge Hubs for the Life Sciences*. Washington, DC: Brookings Institution Press.

Stokes, Donald. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, DC: Brookings Institution Press.

Strickland, Stephen P. 1989. *The Story of the NIH Grants Program*. Lanham, MD: University Press of America.

Van Slyke, C. J. 1946. "New Horizons in Medical Research." *Science* **104**(2711): 559-567.

Walsh, John P., Charlene Cho, and Wesley M. Cohen. 2005. "View from the Bench: Patents and Material Transfers." *Science* **309**(5743): 2002-2003.

Wapner, Jessica. 2013. *The Philadelphia Chromosome: A Genetic Mystery, a Lethal Cancer, and the Improbable Invention of a Life-Saving Treatment*. New York: The Experiment.

Westfall, Richard S. 1985. "Science and Patronage: Galileo and the Telescope." *Isis* **76**(1): 11-30.

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

Wright, Brian D. 1983. "The Economics of Invention Incentives: Patents, Prizes, and Research Contracts." *American Economic Review* **73**(4): 691-707.

Wu, Lingfei, Dashun Wang, and James A. Evans. 2019. "Large Teams Develop and Small Teams Disrupt Science and Technology." *Nature* **566**(7744): 378-382.