Designing Scientific Grants*

Christoph Carnehl† Marco Ottaviani‡ Justus Preusser§

July 2, 2024

Abstract

This paper overviews the economics of scientific grants, focusing on the interplay between the inherent uncertainty in research, researchers' incentives, and grant design. Grants differ from traditional market systems and other science and innovation policy tools, such as prizes and patents. We outline the main economic forces specific to science, noting the limited attention given to grant funding in the economics literature. Using tools from information economics, we identify key incentive problems at various stages of the grant funding process and offer guidance for effective grant design.

In the allocation stage, funders aim to select the highest-merit applications while minimizing evaluation costs. The selection rule, in turn, impacts researchers' incentives to apply and invest in their proposals. In the grant management stage, funders monitor researchers to ensure efficient use of funds. We discuss the advantages and potential pitfalls of (partial) lotteries and emphasize the effectiveness of staged grant design in promoting a productive use of grants.

Beyond these broadly applicable insights, our overview highlights the need for further research on grantmaking. Understudied areas include, at the micro level, the interplay of different grant funding stages, and at the macro level, the interaction of grants with other instruments in the market for science.

*We thank Ben Jones and participants at the NBER EIPE Conference, 2024 for stimulating comments. Marco Ottaviani gratefully acknowledges financial support from MUR-PRIN 2022 (Prot. 2022RL4WMT, "Finanziamento dell'Unione Europea – NextGenerationEU"). Justus Preusser gratefully acknowledges financial support from the European Research Council (HEUROPE 2022 ADG, GA No. 101055295 – InfoEcoScience).

†Department of Economics, Bocconi University, *christoph.carnehl@unibocconi.it*.

‡Department of Economics, Bocconi University, *marco.ottaviani@unibocconi.it*.

§Department of Economics, Bocconi University, *justus.preusser@unibocconi.it*.

Scientific grants—upfront payments to support research in promising yet uncertain areas—are a key source of research funding.[1] In this paper, we explore the economics of grants, with an emphasis on the interplay between the uncertainty inherent in research, incentives and interests of both researchers and funders, and the design of grant schemes. The following historical account from the earliest days of modern science illustrates many of the key issues that prevail until today.

In the summer of 1609, a relatively obscure lecturer of Mathematics at the University of Padua by the name of Galileo Galilei (1609) wrote a letter to the Doge of Venice to present "a new invention of a telescope [...] to bring extraordinary benefit to Your Highness." In actuality, Galileo did not invent but only enhanced the telescope. Nevertheless, he hinted that he would relinquish to the Doge the power to control the further diffusion of telescopes, a dubious claim given the quick diffusion of the tool as a fashionable toy among well-off Venetians. As chronicled by Biagioli (2019), Galileo obtained a salary increase and tenure at the University of Padua.

In the following months, Galileo pointed the telescope at the sky, making a series of astronomical discoveries—among these, the four largest satellites of Jupiter—that marked the dawn of modern science. Galileo hastened to establish priority and published his treatise *Sidereus Nuncius* in March 1610, only two months after his discovery.[2]

Galileo seized the opportunity for a major career advancement and a return to his native Tuscany. He dedicated his treatise to Cosimo de' Medici, the Grand Duke of Tuscany, and named the four satellites of Jupiter that he discovered the Medicean Stars. Galileo (1610) negotiated a prestigious appointment, desiring that "in addition to the title of mathematician His Highness will annex that of philosopher," and sought the financial support that would enable him to focus on research "[...] because giving private lessons and taking scholars as boarders constitute something of an obstacle[.]" Galileo highlighted the prestige his patron would gain, promising "[...] such inventions as no other prince can match, for of these I have a great many and am certain I can find more."

[1] For example, the National Institutes of Health (NIH) in the US had a budget of around 45 billion dollars in 2022 (https://www.nih.gov/about-nih/what-we-do/budget). The Horizon Europe program has a budget of around 95 billion euros for the period 2021-2027 (https://commission.europa.eu/funding-tenders/find-funding/eu-funding-programmes/horizon-europe_en).

[2] Urgency was justified given the rapid diffusion of the telescope. As a matter of fact, Galileo was later involved in a bitter priority dispute with German astronomer Simon Marius, who apparently observed the four satellites of Jupiter one day after Galileo (Pasachoff, 2015).

Galileo also recognized his value as a peer evaluator: "Concerning these inventions which belong to my calling, His Highness may rest assured that he will not be wasting his money on them, as perhaps he has done at other times in great quantity, nor will he miss out on any that are useful and good which have been proposed to him by other men."

Moved by this eloquent appeal, Cosimo appointed Galileo chief mathematician and philosopher to the Grand Duchy of Tuscany in the summer of 1610. Simultaneously, Galileo obtained the position of head mathematician of the University of Pisa, carrying no obligation to reside or teach there—essentially a full teaching buyout.

Although four hundred years have passed since the time of Galileo, his entrepreneurial efforts to secure funding and recognition seem to perfectly anticipate the labors of modern researchers. Like Galileo, researchers appeal to funders by demonstrating a record of successes while promising impactful new discoveries. Like the Doge and the Grand Duke, funders want to channel resources towards talented researchers and rely on researchers' expertise to evaluate funding opportunities. And like Galileo, researchers are under pressure to publish quickly and establish priority.

Unlike traditional market systems widely analyzed by economists, scientific grants are characterized by upfront payments without contractible goals, giving researchers significant autonomy in using these resources. This flexibility is essential due to the nature of grant-supported research, which is often open-ended and uncertain. For example, it is typically impossible to define the value or outcome of research in advance, given its inherently open-ended and exploratory nature. Neither funding agencies nor researchers can anticipate whether years of research will ever lead to a significant discovery, whether the scientific community will realize the significance, or whether the discovery will have an impact beyond the scientific community. Driven by passion, intellectual curiosity, and the hope for scientific recognition, investigators conduct "blue-sky" research to solve open puzzles and challenge conventional wisdom. More than perhaps any other source of funding, grants thus leverage researchers' self-motivation for discovery and recognition within the community.

Compared to other funding instruments, such as patents and research prizes, grants seem to have received considerably less attention in the economics literature. We are confronted with at least two questions.

First, what economic reasons are there for funding science using grants instead of patents or research prizes? Figure 1 offers a schematic view of science funding. Grants

are awarded before researchers begin working on their projects. Upon successfully completing research projects, researchers stand to gain intellectual satisfaction, recognition within the community, and patents and prizes. Existing work has illuminated the impact of patents and prizes on researchers' incentives. How do these incentives interact with the incentives provided by scientific grants? What are the relative strengths and weaknesses of these instruments?

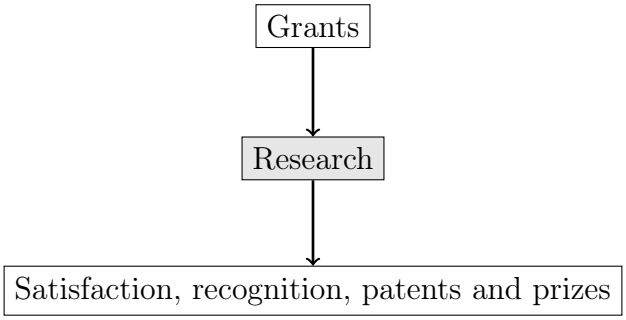


Figure 1: A schematic view of science funding.

Second, if one uses grants, what should they look like, and how should they be allocated? Figure 2 gives a schematic view of the grant funding process. The funding process involves a number of stages where either researchers (in the shaded boxes) or grantmakers (in the unshaded boxes) make decisions. How does prospective evaluation impact the incentives of researchers to invest in the intrinsic quality or persuasiveness of their proposals? What measures can grantmakers take to stimulate or mitigate these investments? How can retrospective evaluation be used to effectively screen applicants and manage grantees?

We draw from information economics and mechanism design to gain insight into these issues. Spread across a diverse set of papers, the literature has developed several important results that shed light on individual pieces from Figures 1 and 2. However, in our view, more remains to be done to develop a unifying theory on the relationship between different pieces, clarifying the interaction of different stages of grant funding (Figure 2) and the interaction between different stages of science funding (Figure 1).

Our work complements recent overviews by Price (2019) and Azoulay and Li (2021). Price (2019) explains the details of grant allocation at the National Institutes of Health (NIH) and offers a rebuttal against common critiques of grants. Azoulay and Li (2021) review the empirical economics literature on grants and discuss various aspects of practical grant design.

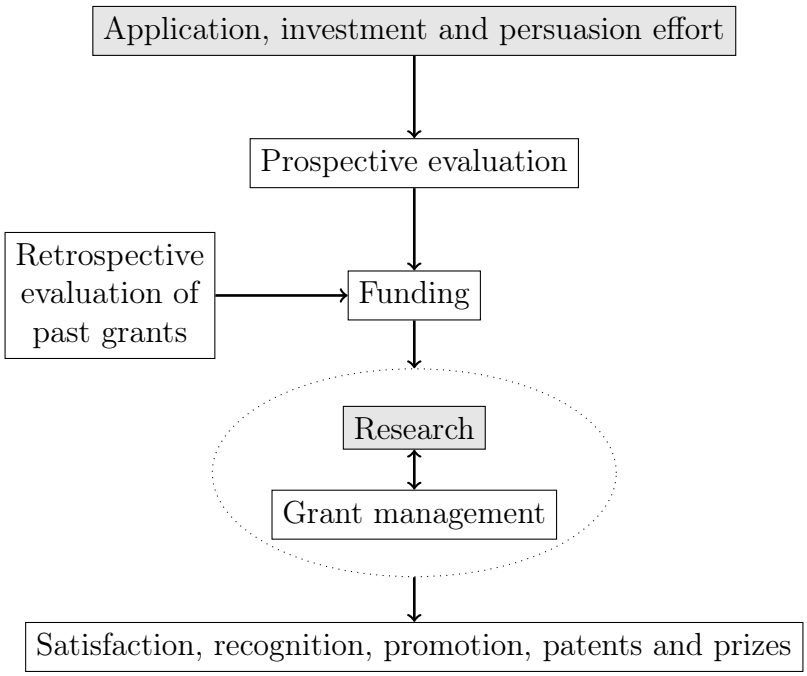


Figure 2: A schematic view of grant funding.

The paper proceeds as follows. In the next section, we briefly discuss the economic justification for science funding, and we delineate some of the situations in which grants appear more suitable than patents and research prizes. The remainder of the paper focuses on grants. In Section 2, we lay out the basic incentive problems that arise in grant allocation. In Sections 3 to 5, we then discuss in more detail the economics of the application and allocation processes, as well as the post-award management stage. Finally, in Section 6, we ask how the supply of grants impacts the direction of research.

1 Why and how to fund science?

Scientific research is central to the economic well-being of modern societies since it creates valuable new knowledge. As explained by Arrow (1962), however, there are several reasons why one may think that markets fail to incentivize an efficient production of knowledge. First, while scientific research consumes large amounts of resources, its output, knowledge, is non-rival since it can be reproduced at (almost) no cost. Second, knowledge production creates positive externalities by inspiring

follow-up research. Third, there is substantial uncertainty about the nature and value of research output. It is difficult to insure against this uncertainty since one needs specialist knowledge to evaluate both a project's ex-ante prospects and, at the ex-post stage, the work that was actually carried out.

For these reasons, there is an economic justification for government interventions that adjust incentives in the market for knowledge production. Governments around the world recognize the importance of science policy for the prosperity of societies in general, but also as a strategic investment in national competitiveness and economic security (Dasgupta and David, 1994). Jones (2021) summarizes a large body of evidence that points to the large social gain from subsidizing science and innovation.

1.1 Incentives in the production of knowledge

Even accepting the need for policies to encourage knowledge production, it is not at all clear what these policies should be. In practice, researchers' incentives are shaped by various monetary and non-monetary instruments, as illustrated in Figure 1.

Grants act at the ex-ante stage and provide *push incentives*. Specifically, grants provide researchers with resources to carry out projects (in the form of specialized equipment and access to data, for example) and reduce the costs of conducting research (by funding personnel, for example). Thus, grants *push* researchers into a broad direction while leaving them with substantial leeway to choose the exact topic and method. Since grants are awarded before any knowledge is produced, the allocation of grants relies on the costly ex-ante evaluation of proposals and researchers.

Patents and prizes offer rewards only for completed projects. Thus, patents and prizes *pull* researchers towards questions that the market or the prize-maker values. Accordingly, costly evaluation happens ex-post when inventors apply for patents (Matcham and Schankerman, 2023) or the prize committee chooses a winner.

In addition to these monetary incentives, there are non-monetary incentives at work, stemming from intellectual satisfaction and recognition in the scientific community from establishing priority. Merton (1957) emphasizes the importance of the priority system for researchers' incentives. Hill and Stein (2023) find that narrowly failing to establish priority has a negative impact on publication and career outcomes for structural biologists.

1.2 Why fund science using grants?

We next delineate situations in which grants seem particularly useful. There appear to be no formal economic analyses comparing grants to other funding instruments. Earlier work by Gallini and Scotchmer (2002) and Wright (1983) views grants as special fixed-price contracts and offers only a stylized analysis.

To begin with, we note that the non-monetary incentives stemming from priority and citations may be insufficient for encouraging an efficient production of knowledge. The priority system and citations reward researchers proportionally to the impact of their work, which would seem to encourage work on important problems. Alas, the system is imperfect. The pressure to publish can have a detrimental impact. When several researchers work on the same problem, there is a race to be the first to publish. Hill and Stein (2024) provide evidence that structural biology projects with higher recognition potential are completed faster and are of lower quality. Further, there is evidence of publication bias against null results (see Andrews and Kasy, 2019, for example). The importance of citation measures in quantifying scientific impact and in promotion decisions can induce researchers to, possibly inefficiently, prioritize research fields that generate higher citation counts over others (see Olszewski, 2020, for example). Thus, even though non-monetary incentives surely play an important role, it is worth examining the role of monetary incentives.

Compared to patents and research prizes, grants have several strengths. First and most straightforwardly, grants address financial constraints and allow funders to target specific people and institutions. Scientific apprenticeships, laboratory equipment, and scientific infrastructure are typically financed via grants (Azoulay and Li, 2021; Price, 2019). Thus, grants can direct resources towards minorities or compensate endowment differences across universities to level the playing field.[3] Patents and prizes, offering rewards only for completed projects, arguably do less than grants to alleviate financial constraints. While prizes can be targeted towards specific people and institutions, patents are typically untargeted.

Second, grants may be useful for aligning market values with social values. A mismatch between the two values may arise when market participants are unwilling or unable to pay for research output; Azoulay and Li (2021) and Price (2019) give the example of medical treatments whose social value far exceeds any individual's

[3]Yang et al. (2022) show that more diversity in teams can lead to better research outcomes.

willingness to pay. Castillo et al. (2021) indicate an enormous gap between the social and commercial incentives to expand vaccine capacity during the COVID-19 pandemic. A mismatch between social and market values diminishes the appeal of funding instruments that, like patents, create incentives by linking the rewards from research to market returns. Bryan and Lemus (2017) provide theoretical support for the idea that patents fall short of guiding researchers to the right topics. In their framework, researchers choose both the intensity and the direction of their research. They show that, while patents can induce the efficient intensity of research, patents do not guarantee directional efficiency. One could conjecture that, here, grantmakers can step in by committing to identifying and supporting socially valuable topics.

Third, grants help address the uncertainty inherent in research (see, for example, Franzoni and Stephan, 2023, for a discussion of uncertainty in research). Fundamental research often involves high uncertainty about the value of discoveries. Further, the commercial potential of discoveries may reveal themselves only much later.[4] From theoretical work by Aghion et al. (2008), we can glean some intuition for the impact of this uncertainty on funding design. In their framework, researchers are willing to forgo high salaries in exchange for freedom of academic pursuit, consistent with empirical evidence by Stern (2004). Aghion et al. (2008) argue that, in the early stages of research, which is when the commercial value of projects is most uncertain, it is efficient to leave scientists the freedom to decide which questions to pursue, even "if this entails some probability of the scientists wandering off in other directions." Interpreting grants as providing unconditional funding for researchers, this provides a rationale for funding fundamental research via grants.

Fourth, grants may be most suitable when it is difficult to spell out what exactly constitutes a successful research project. In some cases, it is roughly clear what success means. For example, in 1908, the Wolfskehl Prize was announced for the first person to prove Fermat's Last Theorem. In 1997, the prize was awarded to Andrew Wiles after the scientific community deemed his proof complete. However, the outcome and the value of an undertaking are often highly uncertain. In this case, an upfront payment in the form of a grant may constitute a more compelling incentive to tackle a problem than, say, a prize for attaining a nebulously specified goal.

[4]See Azoulay et al. (2019b), Bryan and Williams (2021), and Li et al. (2017) for some evidence that grant-funded projects indeed lay the foundation for patented inventions.

2 Challenges in grant allocation

Despite the merits of grants via-à-vis patents and prizes, grant allocation is not without its own challenges and pitfalls. How can grantmakers identify the most meritorious applications? How can grantmakers be certain that, as Gallini and Scotchmer (2002) put it, grantees do not "take the money and run?" The efficiency of the allocation depends on how well grantmakers can screen applications for their merit ex ante and on whether funded researchers work hard to execute their proposals. Additionally, one would like to achieve all this while economizing on the costs of preparing and reviewing proposals, monitoring funded researchers, etc.

A central premise of modern economics is that institutions are marred by the simultaneous presence of *misaligned interests* and *asymmetric information*. In the context of a grantmaker deciding whom to fund, it is clear that there are *misaligned interests*: the grantmaker wishes to support the researchers with the "best" project, whereas each researcher may primarily care about their own funding. Asymmetric information is classified into *hidden information* and *hidden action*.

(1) Hidden information refers to the private knowledge of one party that is relevant for the payoffs or strategic incentives of other parties. For example, on the one hand, researchers may have better information than grantmakers about the intrinsic quality of their proposals, but on the other hand, researchers may be uncertain about what exactly grantmakers hope to find in a proposal.

(2) Hidden actions refer to behavior by one party that others cannot observe or control. For example, grantmakers may be unable to enforce exactly how funded researchers use their grants.

Both forms of asymmetric information seem particularly pertinent in research because research topics and practices are highly specific and require specialist knowledge to evaluate. The simultaneous presence of misaligned interests and asymmetric information gives rise to incentive problems.

Prospective and retrospective evaluation are the basic instruments that grantmakers use to mitigate incentive issues. Prospective evaluation (through referee panels, for example) reduces informational asymmetries prior to the allocation. Retrospective evaluation (through publication metrics, for example) rewards researchers for their output, thereby aligning researchers' incentives with those of the grantmaker and mitigating the hidden action problem. Retrospective evaluation also has a prospective

role, as we explain later.

Figure 3 illustrates schematically the stages of grant funding. The shaded boxes indicate stages where applicants make choices; the unshaded solid boxes indicate stages where the grantmaker makes a choice; the dashed boxes indicate an impact of outside factors. When preparing an application, researchers choose how much effort to invest into the intrinsic quality and persuasiveness of their proposals. When deciding who to fund, the grantmaker both prospectively evaluates the submitted proposals and retrospectively evaluates the researchers' records. Unsuccessful researchers may be temporarily excluded from reapplying for the same grant. For successful researchers, the interaction with the grantmaker continues; the researchers decide how to use their funds, while the grantmaker monitors the researchers' output and decides whether to, say, terminate the grant prematurely. Finally, successful projects reward researchers with accolades in the scientific community and, possibly, monetary rewards in the form of patents and prizes.

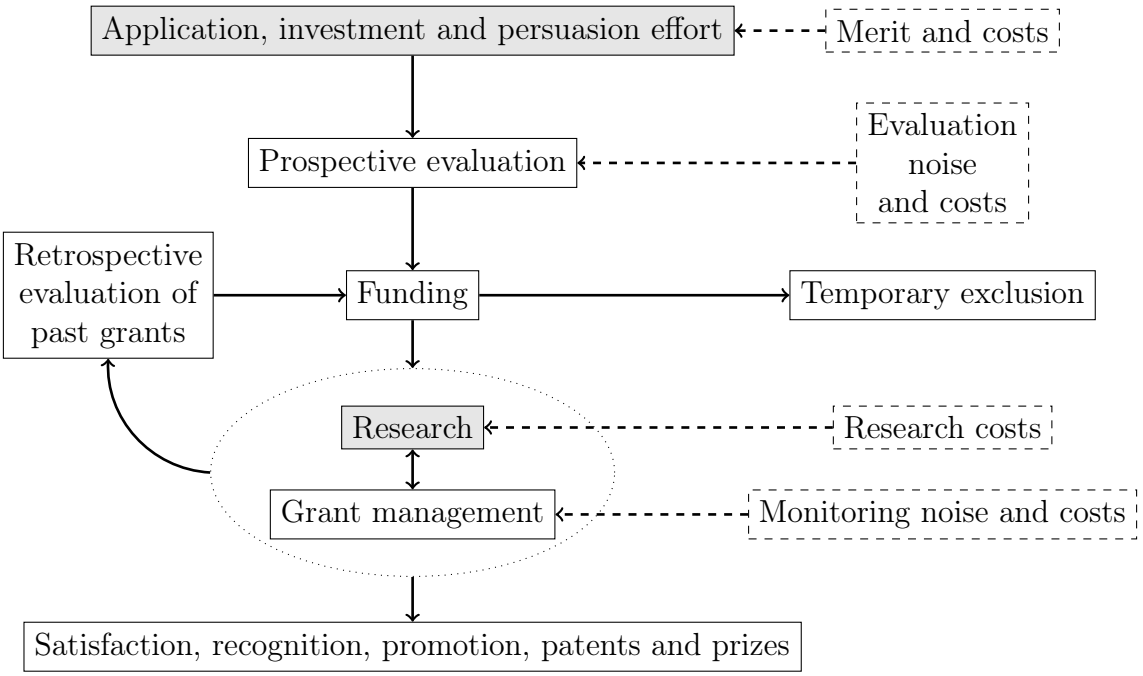


Figure 3: A schematic view of grant funding.

In the next section, we analyze in more detail how the forces indicated in Figure 3 shape the optimal allocation and use of grants. In Section 3, we study the application process, emphasizing the impact of evaluation noise and application costs. In Section 4,

we focus on the allocation rule itself. Next, in Section 5 we turn to the role of retrospective evaluation and post-award management. Finally, in Section 6 we consider how the grantmaker's choice of what to fund impacts the direction of research.

3 The application process

Grant programs thoroughly evaluate proposals via peer review. The NIH, for example, is required by law to do so (Price, 2019). Only the applications deemed to be most promising are funded. For an indication of how competitive the process is, consider that the fraction of funded applications for "Advanced Grants" at the European Research Council (ERC) is around 10%. Preparing an application is a time-consuming and costly activity;[5] researchers are more willing to apply when they are more optimistic about obtaining funding. The costly ordeal of preparing an application thus helps grantmakers screen applicants through self-selection. Conversely, application rates respond to changes in application costs and the evaluation process.

Adda and Ottaviani (2023) analyze the impact of prospective evaluation on the self-selection of applicants. The baseline version of their model can be understood in terms of demand and supply for grants. A more detailed description of their baseline model is in Model Box 1, further below.

On the demand side, researchers differ in their intrinsic and privately known "merit." To give a concrete example, suppose that merit corresponds to the number of additional papers a researcher can publish if awarded the grant. Researchers enjoy a benefit if funded, but bear application costs.

On the supply side, the grantmaker can only fund a fraction of the applicants and would like to assign grants to the most meritorious applicants. However, the grantmaker only observes a noisy signal about each applicant's merit (via peer review, for example). From the grantmaker's perspective, a higher signal indicates a higher merit of the applicant. Therefore, the grantmaker funds the applicants that obtain a signal above an *acceptance threshold*; the threshold is chosen to exhaust the grantmaker's budget.

Anticipating this allocation rule, a researcher applies only if they are sufficiently

[5]From surveys of astronomers, Hippel and Hippel (2015) report an average time of 116 hours spent preparing one application. Myers (2024), also using survey data, indicates researchers spend 7 hours per week preparing grant applications.

optimistic about being evaluated positively; that is, about generating a signal above the acceptance threshold. Therefore, there is an *application threshold* (distinct from the acceptance threshold) such that only researchers with merit above the application threshold will apply.

Building on this characterization, Adda and Ottaviani (2023) shed light on the effect of changes in the grantmaker's budget and evaluation procedure. First, an increase in the budget of the grantmaker naturally increases the incentive to apply, thereby reducing the average merit of applicants. The effect of a budget increase on the success rate—the ratio of funded applicants to the total number of applicants—is more subtle and depends on the distribution of merit across researchers. On the one hand, a budget increase incentivizes more applications (as just noted), pushing the success rate down. On the other hand, a budget increase allows for more applicants to be funded, pushing the success rate up. Thus, the success rate can either increase or decrease. In particular, it decreases if the elasticity of applications with respect to the budget is greater than one.[6] This appears in line with evidence from the 2009 increase in the budget available for research grants in the US due to Obama's Stimulus Package. Grant applications increased more than the budget, thus resulting in a reduction in the fraction of successful applicants (Stephan, 2012, p. 145).

A key insight of Adda and Ottaviani (2023) pertains to the impact of evaluation noise on application incentives. The noise could reflect how carefully evaluators read proposals or whether a funding lottery is used for applicants at the cusp. Adda and Ottaviani (2023) show that a noisier evaluation procedure increases the incentives to apply and thus reduces the self-selection of applicants. Intuitively, as the evaluation becomes noisier, the probability of succeeding in obtaining a grant becomes less responsive to merit, encouraging more low-merit applications (we sketch the logic in more detail below). This result indicates how the details of the allocation rule impact the self-selection of applicants. We provide a more detailed derivation of the result in Model Box 1.

[6]Technically, this case arises when the distribution of researchers' merits follows a distribution with thicker tails than exponential. This assumption is in line with an early observation by Lotka (1926) that researchers' productivity in terms of publications follows a power law.

Model Box 1: Grant allocation and evaluation noise

In the following, we set up the single-field model of Adda and Ottaviani (2023) and derive the optimal application behavior and grant allocation rule. Building on the constructed equilibrium, we illustrate how equilibrium outcomes vary with evaluation noise. The grantmaker's budget is denoted by B. For a researcher, applying comes at a cost c, and the value of winning the grant is normalizes to 1. Let θ denote the privately known merit of a researcher. From the grantmaker's perspective, merit is a random variable whose cumulative distribution function (CDF) we denote by G. The grantmaker observes a noisy signal x about merit. Conditional on merit θ, the signal x is a random variable with CDF F_θ, where higher signal realizations indicate higher merit.[a]

Since signals are informative about merit, the grantmaker optimally sets an acceptance threshold $\hat{x}$ such that all applicants with a signal above $\hat{x}$ are funded. From the perspective of a researcher with merit θ, the winning probability, denoted $W(\theta;\hat{x})$, is thus given by $W(\theta;\hat{x}) = 1 - F_\theta(\hat{x})$. Such a researcher applies if and only if the winning probability exceeds the application costs, meaning $W(\theta;\hat{x}) \geq c$. We denote the application threshold by $\hat{\theta}$. This threshold depends on $\hat{x}$ and coincides with the marginal researcher who is indifferent to applying, meaning $W(\hat{\theta};\hat{x}) = c$ holds. All researchers with merit $\theta \geq \hat{\theta}$ apply.

In equilibrium, the acceptance threshold $\hat{x}$ is such that the number of awarded grants equals the grantmaker's budget B. As each researcher of merit $\theta \geq \hat{\theta}$ wins with probability $W(\theta;\hat{x})$, the threshold $\hat{x}$ solves $B = \int_{\hat{\theta}}^{\infty} W(\theta;\hat{x})\, dG(\theta)$.

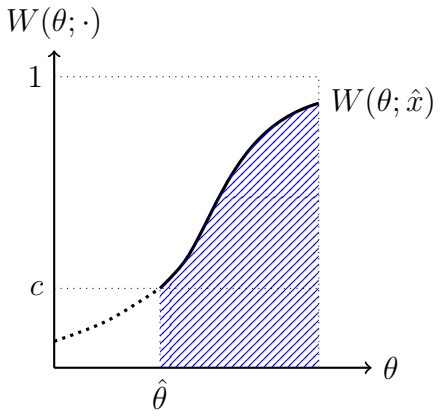


Figure 4: Baseline Equilibrium

Figure 4 depicts this situation for a given threshold $\hat{x}$ assuming that θ is uniformly distributed. The blue-shaded area in the figure thus represents the mass of awarded grants.

Consider now a noisier evaluation process, represented by a less accurate signal distribution $\tilde{F}_\theta(x)$.[b] Noisier evaluation corresponds to less meritocratic acceptance decisions; the probability of winning the grant is *less steeply* increasing in merit. Pictorially, the function W changes to a flatter function $\tilde{W}$. Suppose for a moment that the new acceptance threshold $\tilde{x}$ is such that the same merit type $\hat{\theta}$ is indifferent; that is, $\tilde{W}(\hat{\theta};\tilde{x}) = 1 - \tilde{F}_{\hat{\theta}}(\tilde{x}) = c$. The left panel of Figure 5 depicts this scenario. Since the new winning probability is flatter, the grantmaker does not exhaust the budget. The red-shaded area in the left panel of Figure 5 depicts the unspent budget.

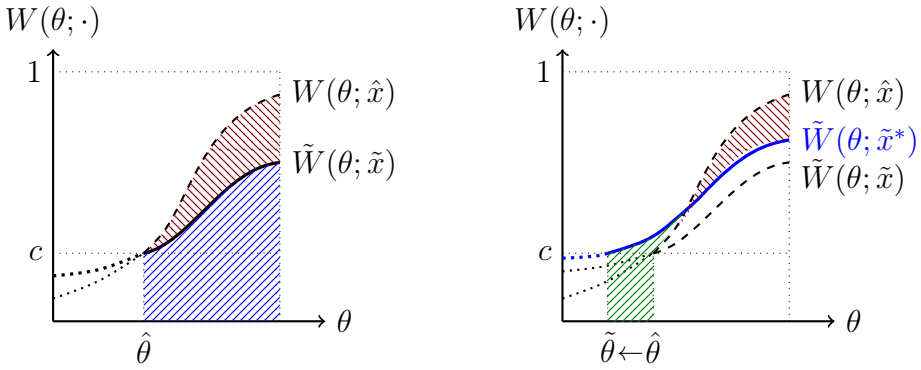


Figure 5: Equilibrium with Noisier Evaluation

To exhaust the budget, the grantmaker has to encourage more applications. That is, the grantmaker decreases the acceptance threshold to a point $\tilde{x}^*$, thereby raising the winning probability $\tilde{W}$ for all merit types. The new marginal researcher lies at a point $\tilde{\theta}$ below $\hat{\theta}$. In the right panel of Figure 5, the green area indicates the volume of additional budget that is spent by decreasing the acceptance threshold to $\tilde{x}^*$; the red and green areas are equal. Thus, in the new equilibrium under noisier evaluation, there are more applications and, therefore, the average merit of applicants is lower.

[a]The distribution F_θ is assumed to possess a continuously differentiable and strictly positive density. Moreover, the density is strictly log-supermodular in signal and merit, meaning the monotone likelihood ratio property holds; higher signals indicate high merit.

[b]Technically, Adda and Ottaviani (2023) consider a decrease with respect to the notion of accuracy introduced by Lehmann (1988).

The apportionment of budget across different fields is more delicate yet. Funding organizations grapple with the incentives of panel members to favor research in their own fields. Panel members have an incentive to inflate the scores of applications in their field in an attempt to secure a larger fraction of the total budget available for distribution to all panels. To counteract this incentive, some of the world's largest research funding organizations have been allocating funds depending on the relative ranking of applications received in each panel, thus apportioning the overall budget according to a mechanical formula. For example, the payline system adopted in 1988 by NIH institutes equalizes the success rate across panels. Since its inception in 2007, the ERC adopted an equivalent scheme by splitting its budget in proportion to the funds requested by applicants in each panel, thus automatically equalizing the success rate across panels.

At first glance, proportional apportionment seems fair and balanced. In the framework of Adda and Ottaviani (2023), the system indeed performs well if fields are relatively similar in terms of the noise in the evaluation signal. However, if fields are heterogeneous in terms of noise, the performance deteriorates. Noisier fields attract more applications (as argued above), which, under proportional apportionment, leads to a proportional increase in the budget; this budget increase, in turn, induces a further increase in applications. To see how adverse this effect can be, consider as an extreme the case of a field with perfect evaluation (meaning all applicants correctly anticipate whether they will succeed). Applicants who anticipate failing do not apply, and thereby reduce the funding for other applicants from the same field. Whenever the success rate is less than 100%, the process will continue until, in equilibrium, no applications at all will be submitted from this field.

Beyond this stark illustration, Adda and Ottaviani (2023) show that under a general class of apportionment rules, including the proportional rule, a reduction of noise in a field leads to a reduction in applications in that field and an increase in applications in all other fields.

Adda and Ottaviani (2023) also present a broad empirical confirmation of their predictions by exploiting a natural experiment. A 2014 reform changed the proportional budget allocation rule at the ERC. After the reform, a panel's budget depended not only on applications in its own domain (e.g., life sciences) but also on the number of applications in the other domains (e.g., social sciences and humanities).

Figure 6 reports the percentage change in budget for each ERC panel after the

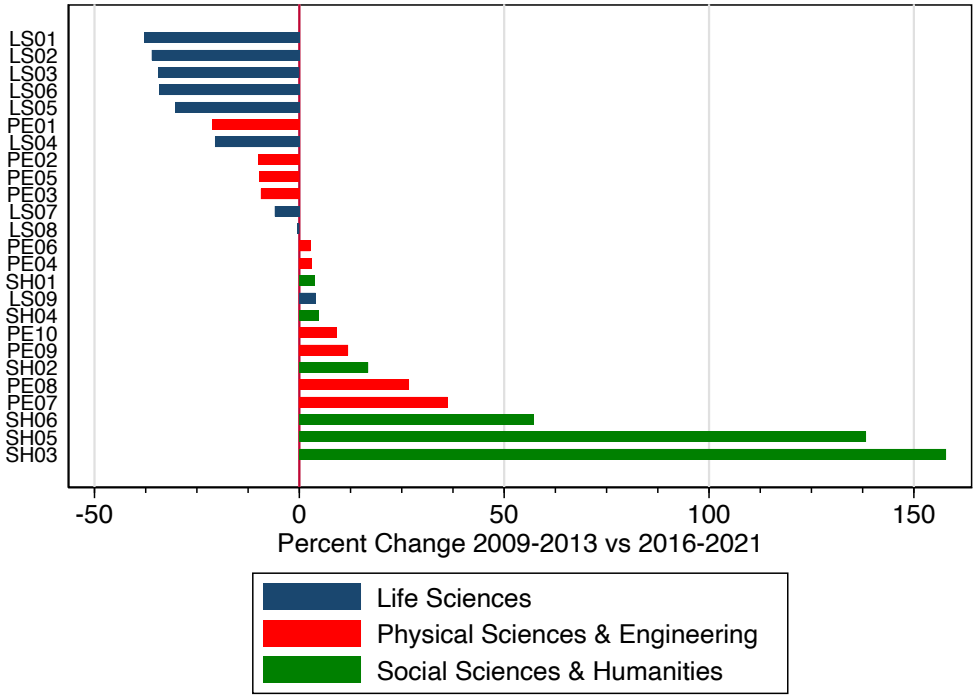


Figure 6: The percent change in the budget of each ERC panel after the reform.

reform in the budget allocation rule, compared to pre-reform levels. Panels in the social sciences and more applied areas of physical sciences and engineering received increased budgets, primarily at the expense of panels in the life sciences and more foundational areas of physical sciences and engineering.

Figure 7 reports a measure of consensus (Gwet's inter-rater agreement) among reviewers for grant proposals submitted to the Research Council of Norway, with abstracts similar to those submitted to the corresponding ERC panels. In line with theory, Adda and Ottaviani (2023) find that on average panels with less agreement, and thus more evaluation noise, gained budget after the reform.

In terms of efficiency, it is undesirable that noisier fields attract more applications and more funding. After all, in noisy fields, it is difficult to allocate efficiently, and hence, all else equal, these fields should be endowed with little funding. An important challenge is to identify rules that are both efficient and politically feasible.

A second important challenge is to reduce the application and evaluation costs inherent to this system. We briefly discuss *temporary exclusion* with this challenge in mind. By temporary exclusion, we mean that unsuccessful applicants are debarred

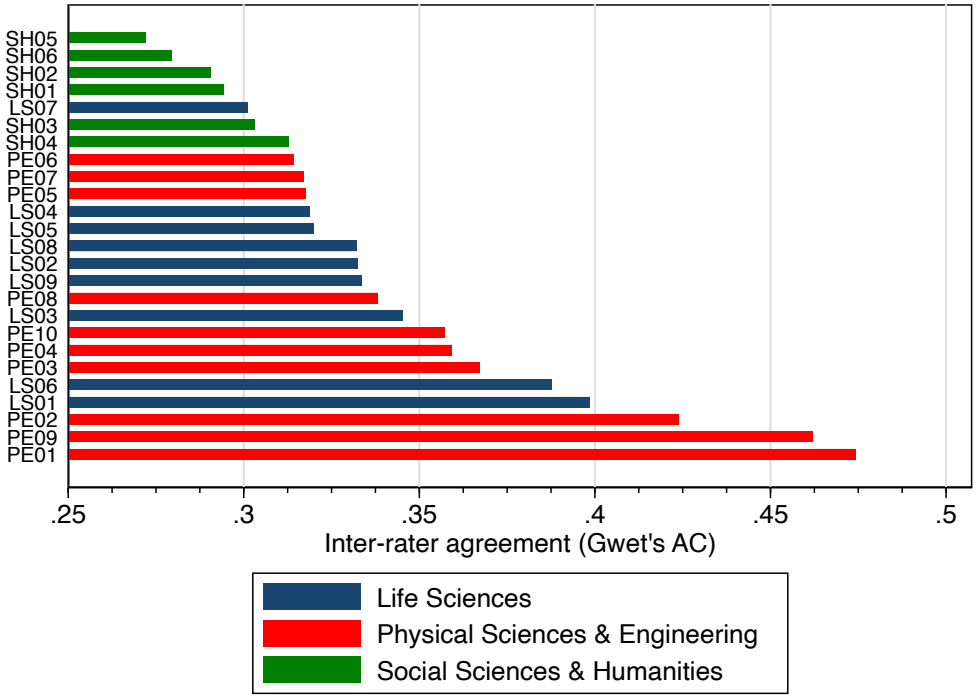


Figure 7: The inter-rater agreement by panel in evaluating RCN funding applications.

from reapplying for some period of time. At the ERC, for example, applicants to Starting or Consolidator Grants who are rejected in the early stages of the evaluation process may not reapply to the same grant for two subsequent years (ERC, 2023, p. 9).[7] The aim of this restriction is "to allow unsuccessful Principal Investigators the time necessary to develop a stronger proposal."

Azrieli (2024) studies temporary exclusion in a dynamic model similar to the model of Adda and Ottaviani (2023). The threat of temporary exclusion deters researchers from applying if they know that their current project is weak, but they expect to have a stronger one in the future. This basic logic suggests temporary exclusion indeed reduces excessive applications. The full picture is more nuanced because of a countervailing force: if one's competitors are less likely to apply, then one's own incentive to apply increases. Azrieli (2024) shows that temporary exclusion can at once increase the overall welfare of researchers and decrease the number of applications. In particular, this result applies when the costs of applying are small relative to the gains from obtaining a grant.

[7] See https://erc.europa.eu/apply-grant/starting-grant.

Further reading Pereyra and Silva (2023) study a related setting where, roughly speaking, the grantmaker allocates grants of heterogeneous quality (for example, they may in the funding duration). As in Adda and Ottaviani (2023), the grantmaker only observes a noisy signal about researchers' privately known types, but researchers bear no application costs. To maximize welfare, the grantmaker lets researchers self-select into one of multiple tracks. In "tougher" tracks, researchers need more positive evaluations to be awarded a high-quality grant.[8]

4 The allocation rule

In the previous section, we focused on the application process, taking a stylized view on how the grantmaker allocates grants to applicants. We now zoom in on optimal allocation rules through the lens of *mechanism design*.

Mechanism design is a subfield of economic theory that studies how to design institutions—*mechanisms*—to achieve a pre-specified goal. Typically, this goal will not be perfectly achievable. For example, a grantmaker may wish to award grants to the most fund-worthy applicants while economizing on the time spent reviewing proposals. The applicants, who know the nuts and bolts of their proposals, have private information about their fund-worthiness that should guide the efficient allocation. However, if the grantmaker would completely save on evaluation costs by never triggering a review, thereby solely relying on what the applicants claim in their proposals, then the applicants would have incentives to overstate their fund-worthiness. Hence, we do not expect that the grantmaker can achieve the first-best outcome of funding the best proposals at zero cost. This impossibility is a consequence of the simultaneous presence of misaligned interests and asymmetric information (and the fact that there are not enough resources for everyone).

When the first-best is out of reach, mechanism design can shed light on achievable *second*-best outcomes: how does a mechanism optimally balance allocative efficiency with evaluation costs? For example, we shall see that optimal mechanisms feature some deliberate randomization in some but not in all environments. Thus, here, the

[8]Pereyra and Silva (2023) do not explicitly cast their model as one of grant allocation. Any faults with the interpretation of their model in the context of grant allocation are due to us. Another important difference to Adda and Ottaviani (2023) is that Pereyra and Silva (2023) cast the problem as one of mechanism design. In particular, the grantmaker has commitment power. We discuss mechanism design in more detail in Section 4.

analysis suggests a practically relevant feature of allocation rules and illuminates which properties of the environment justify this feature from an efficiency perspective. More conceptually, the analysis provides guidance on how grantmakers can benefit from combining certain instruments (such as a costly review process) with self-reports by the applicants.

Notice that we take the objective of the grantmaker as given; the question is how to implement it. We will be silent on what the grantmaker's notion of fund-worthiness should be, or on how the grantmaker's objective should weigh fund-worthiness against the time burden imposed on the reviewers.

Before delving into the details, we emphasize that, unless mentioned otherwise, the models discussed below are cast as general allocation problems. In particular, we do not claim that the papers from which we draw are focused on the specific problem of allocating grants. Any faults with the interpretation of the papers' results in the context of grant allocation lie with us.

4.1 Costly prospective evaluation

How can a grantmaker fund the most promising research while economizing on evaluation costs? We study this question using the model of Ben-Porath et al. (2014).

Consider a grantmaker who has one grant to allocate among a group of applicants. Each applicant wants the grant for themselves and privately knows their merit. For the sake of concreteness, suppose each applicant has some estimate about the number of additional publications they could produce if funded. The grantmaker wishes to maximize this number and has no other use for the funds. If the grantmaker were to ask each applicant to self-report their estimate and then fund the top applicant, the applicants would have an incentive to exaggerate. The grantmaker has an additional instrument at their disposal: they can *verify* individual applicants at a cost. Verifying an applicant reveals the applicant's merit. We interpret the act of verifying as conducting an in-depth review by external evaluators; verification costs represent the time burden this review imposes on the evaluators, capturing the idea that the grantmaker wants to economize on evaluation. We assume here that the applicants are ex-ante symmetric (with respect to the distribution of their merit levels and the costs of verifying them); Ben-Porath et al. (2014) also cover the asymmetric case.

Let us elaborate on the interpretation (Ben-Porath et al., 2014, p. 3803). What

really matters is that verification reveals everything the applicant knows; that is, if the applicant's estimate for the number of additional publications is pinned down by a set of objective facts about the applicant's research that are ex-ante only known to the applicant, then verification reveals those facts. There could well be residual uncertainty about the number of additional publications, and neither the applicant nor the principal can presently resolve this uncertainty.

In this environment, a mechanism specifies which reports the applicants can make to the grantmaker, which applicants are verified depending on those reports, and how the grant is eventually allocated depending on the reports and verification outcomes. To illustrate, here are three examples of mechanisms. First, the grantmaker could randomly select a winner without requesting or verifying reports. Second, the grantmaker could outright verify all applicants and then allocate the grant to the most meritorious one. Third, the grantmaker could approach applicants sequentially, asking for self-declarations about merit, possibly verifying some of them, and stopping when an applicant seems "good enough." Given a mechanism, applicants strategically choose which reports to make to maximize the probability of receiving the grant, forming conjectures about what their competitors might do. The scope of possible mechanisms and outcomes is thus quite complex.

Which mechanisms maximize the expected merit of the funded applicant net of verification costs? Ben-Porath et al. (2014, Theorem 1) show that there is an optimal mechanism of the following form: The grantmaker announces a merit threshold and asks all applicants to self-report their merit. If all reports are below the threshold, the grant is allocated uniformly at random. If, instead, at least one report is above the threshold, then the grantmaker verifies the highest report and, if the report is verified to be true, awards the grant to the corresponding applicant. If an applicant is verified and found to have misreported their merit, then that applicant is *not* allocated the grant.[9] Given these allocation and verification rules, all applicants find it in their best interest to report truthfully; indeed, when untruthfully claiming a merit above

[9]In this case, as long as the misreporting applicant is denied the grant, there is a lot of flexibility regarding how to proceed with the allocation. The reason for this flexibility is that in the described mechanism, the applicants expect that everyone will report truthfully, and therefore, they completely disregard the event in which one of the other applicants is verified to have misreported. One way of proceeding would be to allocate the grant randomly among the remaining agents or not to allocate at all. Another perhaps more reasonable way is to sequentially verify the remaining highest reports until someone is verified to have been truthful, and then award that applicant; in case everyone is verified to have been untruthful, no one is awarded.

the threshold, an applicant anticipates they will be detected. In Model Box 2, we illustrate the one-applicant version of this model and show how to derive the optimal allocation mechanism.

The described optimal mechanism allocates the grant through a lottery when no report is sufficiently promising. This feature speaks to the growing interest in lotteries among real-world funders (Bendiscioli et al., 2022; Heyard et al., 2022). A key argument put forth by Fang and Casadevall (2016) in favor of lotteries is that identifying the best proposals is extremely costly. This concern is roughly reflected in the optimal mechanism. The grantmaker *could* verify all reports and identify the best applicant. However, the grantmaker chooses not to do so, saving on verification costs by randomly allocating the grant when all reported merits are below the threshold. That said, the mechanism identifies and allocates to the best applicant whenever the highest merit is above the threshold. Further, the use of a lottery does depend on the assumption that the applicants are ex-ante symmetric.[10]

Ben-Porath et al. (2014, Section IV) show that for higher verification costs, the optimal mechanism uses a higher threshold; so, there are fewer in-depth reviews, and allocation is more frequently via a lottery. They also show that the threshold increases in response to shifts of the distribution of merit in the sense of first-order or second-order stochastic dominance; that is, roughly speaking, if applicants become more meritorious or less heterogeneous, then the threshold increases, meaning the allocation is more frequently via a lottery.

Model Box 2: Optimal allocation with costly evaluation

Let us consider a single-applicant version of the problem in Ben-Porath et al. (2014). The grantmaker's decision is now between funding the applicant and keeping the funds. The applicant only cares about the funding probability. The grantmaker's payoff from keeping the funds is normalized to 0. Let θ denote the applicant's privately known merit type. The grantmaker's payoff from funding the applicant equals θ, and θ may be greater or less than 0. From the grantmaker's perspective, the merit type θ is a random variable whose cumulative distribution function we denote by F. The verification costs are

[10] In the asymmetric case, there is an optimal mechanism that (for almost all realizations) *deterministically* allocates to a so-called "favored agent" if no self-report passes the threshold (Ben-Porath et al., 2014, Theorem 1).

denoted by c.

For optimality, nothing is lost by focusing on mechanisms of the following form.[a] For each merit type $\hat{\theta}$, the mechanism specifies two probabilities $v(\hat{\theta})$ and $q(\hat{\theta})$, and the mechanism plays out as follows. First, the applicant is asked to report their merit. Given the applicant's report $\hat{\theta}$, the applicant is verified with probability $v(\hat{\theta})$; if the applicant is verified, the applicant is funded for sure if the report was truthful ($\theta = \hat{\theta}$), and with probability 0 if untruthful; if the applicant's report is not verified, the applicant is funded with probability $q(\hat{\theta})$. Moreover, in the mechanism, the applicant always finds it in their best interest to report truthfully; that is, for all θ and $\hat{\theta}$,

$$v(\theta) \cdot 1 + (1 - v(\theta)) \cdot q(\theta) \geq (1 - v(\hat{\theta})) \cdot q(\hat{\theta}) \tag{IC}$$

The left side of this inequality is the overall winning probability for merit-type θ when reporting truthfully; the right side when misreporting $\hat{\theta}$. This inequality is known as *incentive compatibility (IC) for type* θ.

It is convenient to define the overall winning probability as $p(\theta) = v(\theta) \cdot 1 + (1 - v(\theta)) \cdot q(\theta)$. Thus (IC) can be written as $p(\theta) \geq p(\hat{\theta}) - v(\hat{\theta})$. Since this inequality holds for all θ and $\hat{\theta}$, it also holds if we flip their roles, yielding $p(\hat{\theta}) \geq p(\theta) - v(\theta)$. We can now rearrange to obtain

$$v(\theta) \geq p(\theta) - p(\hat{\theta}). \tag{IC*}$$

This inequality says that the probability $v(\theta)$ of verifying θ must be sufficiently high to deter $\hat{\theta}$ from misreporting to θ.

The grantmaker's expected payoff equals

$$\mathbb{E}_{\theta \sim F}[p(\theta) \cdot \theta + (1 - p(\theta)) \cdot 0 - v(\theta) \cdot c].$$

We want to maximize the grantmaker's payoff subject to (IC*). To see how this is done, let $\underline{p}$ denote the winning probability of the worst-off merit types; that is, $\underline{p} = \inf_{\theta \in \mathbb{R}} p(\theta)$. The truth-telling constraint (IC*) is hardest to satisfy for the worst-off merit types; that is, (IC*) holds if and only if $v(\theta) \geq p(\theta) - \underline{p}$ holds for all θ. The grantmaker wishes to save on verification costs and therefore sets

$v(\theta)$ as low as possible, meaning $v(\theta) = p(\theta) - \underline{p}$. Plugging this equation into the grantmaker's expected payoff and rearranging yields

$$\mathbb{E}_{\theta\sim F}[(p(\theta) - \underline{p}) \cdot (\theta - c) + \underline{p} \cdot \theta].$$

This expected payoff has the following interpretation: when the grantmaker wishes to raise $p(\theta)$ beyond the value $\underline{p}$, the grantmaker must simultaneously raise the verification probability $v(\theta)$ since, else, the worst-off merit types would misreport. Thus, the gain from raising $p(\theta)$ effectively equals the *net* merit type $\theta - c$. It follows that, optimally, $p(\theta)$ equals 1 when the net merit type is positive ($\theta - c \geq 0$) and equals 0 when the net merit type is negative.

The optimal value of $\underline{p}$ depends on the parameters. For example, setting $\underline{p} = 0$ is optimal if the prior mean $\mathbb{E}_{\theta\sim F}[\theta]$ is less than 0 (which was the normalized value of keeping the funds), capturing a scenario where the grantmaker is a priori unwilling to fund the applicant.[b]

When $\underline{p}$ equals 0, the optimal mechanism can be described as follows. The applicant reports truthfully. Reports below 0 are not verified and do not lead to funding. Reports $\hat{\theta}$ above c are verified with certainty and lead to funding if and only if the report is truthful.

[a] This step combines a version of the Revelation Principle with optimality considerations.

[b] Let us show this. Suppose $p(\theta)$ equals 1 if $\theta - c \geq 0$, and equals $\underline{p}$ otherwise. We can now write the grantmaker's expected payoff as

$$(1 - \underline{p})\mathbb{E}_{\theta\sim F}[\max(0, \theta - c)] + \underline{p}\mathbb{E}_{\theta\sim F}[\theta].$$

In this expression, the first expectation is positive. The second is negative (by assumption). Thus the grantmaker's expected payoff is decreasing in $\underline{p}$; optimally, $\underline{p}$ equals 0.

In practice, grantmakers may not only rely on the outcome of an in-depth review; they also have access to a relatively cheap but noisy signal, perhaps from a preliminary screening of the applicants. Kattwinkel and Knoepfle (2023) study how to optimally combine such a noisy signal with self-reports and verification in a setting with a single applicant (thus, the decision is between funding the applicant and keeping the funds). Kattwinkel and Knoepfle show that, optimally, the grantmaker reveals the results of the preliminary review to the applicant, and then allows the applicant to submit an appeal to be verified. The applicant is funded if the preliminary review is

sufficiently positive, or if the applicant is verified to have sufficiently high merit upon submitting an appeal. An intriguing feature of this mechanism is that the grantmaker does not benefit from obscuring the result of the preliminary review; the optimal transparent procedure does not sacrifice allocative efficiency. This feature is surprising since, in theory, the principal could try to cross-check the agent's self-report with the preliminary review to deliver stronger incentives for reporting truthfully, potentially letting the grantmaker save on verification costs.

Further reading Ben-Porath et al. (2019, 2023) consider, among other things, nearby settings where the grantmaker cannot verify the applicants' claims; instead, the applicants themselves can provide hard evidence about their merit but must be incentivized to do so. Khalfan (2023) studies a problem where applicants only have a noisy signal about their merit and hence do not know what an in-depth review will reveal. This detail complicates the grantmaker's problem as applicants can now feign ignorance if a review reveals low merit. Epitropou and Vohra (2019) consider a problem where the applicants arrive sequentially. Li and Libgober (2023) consider a repeated interaction with a single applicant.

4.2 Costly ordeals

As highlighted by Adda and Ottaviani (2023), costly ordeals can help grantmakers screen applicants and improve allocative efficiency. But, of course, an ordeal itself is a wasteful activity, raising the question of optimally balancing costly ordeals and allocative efficiency from a welfare perspective. We explore this question in a different framework based on Condorelli (2012). As we highlight below, depending on the parameters of the environment, it may be optimal to allocate efficiently while imposing highly costly ordeals, but it may also be optimal to simply allocate grants at random while imposing no costly ordeals.

In the model, a grantmaker has a fixed set of grants to be distributed among a fixed set of researchers. The grants potentially differ in their desirability (for example, in the amount of funds or the funding duration). There are not enough grants for everyone. Researchers differ in their privately known valuations for funding (we interpret this valuation momentarily). Prior to allocating the grants, the grantmaker asks researchers to report their valuations and then imposes costly ordeals on the

researchers. An ordeal may consist of filling out forms whose length is commensurate with one's claimed valuation but with no informational value on top of one's report. Researchers with higher valuations will have a higher willingness to bear any given ordeal. Of course, researchers can walk away at any time to avoid the ordeal and forgo the chance at funding. The grantmaker seeks to maximize the expected valuation across funded researchers minus the costs of the imposed ordeals, potentially assigning different welfare weights to different researchers.

We interpret a researcher's valuation as capturing, for example, the privately known scientific merit of the researcher's project as well as how desperate the researcher needs funds for purchasing laboratory equipment. In this interpretation, the researcher's willingness to endure an ordeal is increasing in their merit and funding need.

Condorelli (2012) characterizes the following optimal mechanism. The grantmaker computes for each researcher and reported valuation a *priority* (explained momentarily) and then assigns the most desirable grants assortatively to the researchers with the highest priorities (until all grants are assigned). The ordeals are chosen to elicit truthful claims. Researchers with higher claimed valuations bear more costly ordeals to deter exaggerated claims.

A researcher's priority is increasing in their reported valuation, but not necessarily strictly increasing. That is, researchers of different valuations may be assigned the same priority. As described by Condorelli, the exact relationship between valuation and priority depends on the prior distribution of valuations. We highlight two cases.

First, suppose each researcher's valuation distribution has a *weakly increasing* hazard rate. In this case, each researcher's priority is constant in the reported valuation; a researcher's priority depends only on their expected valuation and their welfare weight in the grantmaker's objective function. The described optimal mechanism thus assigns the grants *without* eliciting reports or imposing ordeals. If researchers are symmetric (in terms of valuation distributions and welfare weights), the mechanism simplifies to a lottery.

Second, suppose each researcher's valuation distribution has a *strictly decreasing* hazard rate. In this case, each researcher's priority is strictly increasing in their valuation. If researchers are symmetric, the described optimal mechanism thus allocates the best grants to the highest valuation researchers. The mechanism maximizes allocative efficiency but relies on costly ordeals to elicit the researchers' private information.

Further reading See work by Chakravarty and Kaplan (2013), Hartline and Roughgarden (2008), and Yoon (2011) for similar analyses. Dworczak (2023) provides a sufficient condition for optimal mechanisms to involve costly ordeals. Dworczak's model is mathematically similar to Condorelli's, but the interpretations differ somewhat; see Dworczak (2023, Section 2.1) for elaboration.

4.3 Peer selection

Next, we discuss a setting in which the grant applicants and the reviewers are the same group of people.

On an abstract level, this setting fits the problem of distributing resources within the scientific community. There is a finite amount of resources to allocate. To allocate these efficiently, grantmakers would benefit from knowing how each scientist evaluates their peers. Of course, there is no single grantmaker who allocates all resources, and not all scientists simultaneously apply for a grant, but the model serves as a useful abstraction to shed light on an economic tension (explained momentarily).

On a more concrete level, it has been suggested to *require* applicants to evaluate their competitors' proposals. Merrifield and Saari (2009) suggest such a procedure for allocating telescope time to alleviate the burden that a conventional process places on external reviewers. A 2014 NSF pilot later used such a procedure to allocate grants; the Gemini Observatory to allocate time on its telescope in Hawaii (Mervis, 2014). The 2016 Neural Information Processing Systems conference, a top-tier machine learning conference, asked authors to volunteer as reviewers to handle the enormous number of submissions (Shah et al., 2018).

Here, we focus on one particular issue that appears relevant when applicants review their competitors: dishonestly exaggerating one's merits and others' faults. See, for example, the laboratory experiment of Balietti et al. (2016) or the field experiment of Hussam et al. (2022) for evidence that individuals indeed misreport when evaluating their peers (albeit not in the context of scientific grant funding).

We consider a stylized setting, borrowing from work by Alon et al. (2011). Suppose each applicant has an honest private opinion about who they would nominate as the most deserving recipient of the grant. The grantmaker wishes to allocate to the applicant with the most nominations. At the same time, the grantmaker wishes to use a mechanism under which no applicant can influence their own chance of winning

by misreporting their nomination. In the literature, such a rule is called *impartial*.[11]

The key economic tension in this problem stems from the grantmaker's desire to elicit a relative ranking of the applicants. Applicants improve their rank not only by appraising themselves but also by diminishing their peers. Even if self-nominations are prohibited, an applicant may be able to improve their rank by claiming that all others are unworthy of the grant.

Here is one example of an impartial mechanism. First, before eliciting any nominations, the grantmaker randomly partitions the pool of applicants into two groups—"evaluators" and "candidates". Second, the grantmaker counts only the nominations of the evaluators to pick a winner among the candidates.

Under this partition mechanism, evaluators know that they are excluded from winning (and hence cannot improve their chances by misreporting), and candidates know that their reports are not counted. Instead of randomly splitting applicants into two groups to award one grant, one could also split the grant into two parts and have the evaluators for the first part be the candidates for the second part and vice versa.

The partition mechanism is intuitive and takes to heart a central tenet of peer review—evaluators should have no personal stake in the funding decision—but the mechanism is not flawless. The partition mechanism yields a bad outcome if all applicants with few nominations are randomly assigned the role of candidate.

Fischer and Klimm (2014) show that the grantmaker can improve on the partition mechanism by splitting the applicants into more than two groups (the procedure for aggregating their nominations becomes more intricate, though). Fischer and Klimm propose a mechanism that guarantees to select an applicant who receives at least half as many approvals as the applicant with the most approvals. The fraction of one-half is the best theoretical guarantee (Alon et al., 2011; Fischer and Klimm, 2014). Niemeyer and Preusser (2023) provide conditions under which the partition mechanism is approximately optimal among all impartial mechanisms when there are many applicants, albeit under a different notion of optimality.[12]

[11] Some papers instead use the terms *strategyproofness* or *dominant-strategy incentive-compatibility* instead of *impartiality*. These notions all coincide under the assumption that each applicant seeks to maximize their own chance of winning and is otherwise indifferent to who wins.

[12] Alon et al. (2011) and Fischer and Klimm (2014) evaluate a rule using a worst-case criterion: across all possible nomination realizations, in the worst case, what is the ratio between the nominations received by the chosen applicant and the most nominations received by any applicant? Niemeyer and Preusser (2023) instead consider the number of nominations of the chosen applicant, taking expectations with respect to the grantmaker's belief about the realized nominations.

Further reading A large number of papers build on the framework of Alon et al. (2011). For example, Kurokawa et al. (2015) impose the additional constraint that each applicant can only review a limited number of other proposals. De Clippel et al. (2008), Holzman and Moulin (2013), and Mackenzie (2015), among others, instead study which other desirable properties are compatible with impartiality. Olckers and Walsh (2023) survey the literature on impartial mechanisms. In related settings, Bloch et al. (2023), Kattwinkel (2020), and Kattwinkel et al. (2022) consider mechanisms that are not necessarily impartial but which nevertheless provide the applicants with incentives for truthful reports;[13] for example, they allow for mechanisms where the grantmaker cross-checks reports across applicants or with external reviewers.

4.4 Incentivizing investment

Up to this point, we treated each applicant's merit as exogenous; applicants do not control the quality of their proposals. However, we would think that in practice researchers adjust their proposals in response to, say, the fierceness of their competition or the grantmaker's evaluation criteria.

In this section, we discuss how the grantmaker's mechanism incentivizes productive investment, based on work by Augias and Perez-Richet (2023). The headline insight is that the grantmaker does *not* benefit from allocating randomly. Moreover, the optimal evaluation is noisy for low-merit proposals but exact for high-merit ones.

In the model, a unit mass of grants can be assigned to a unit mass of applicants. Applicants are heterogeneous with respect to their initial merit level. Crucially, applicants can invest in their merit, incurring a cost that increases in the distance between the initial and final merit. Applicants care about the probability of receiving funding minus the investment costs. The grantmaker observes each applicant's final merit (but neither initial merit nor investment) and then decides whom to give a grant. In the baseline model, the grantmaker has enough resources to fund everyone, but would only like to fund those whose final merit exceeds some threshold; let us call this the grantmaker's preferred threshold.

Augias and Perez-Richet (2023, Theorem 1) provide a condition on the distribution

[13] In technical terms, these authors consider *Bayesian incentive-compatible* mechanisms. As noted in footnote 11, impartiality coincides with the stronger notion of dominant-strategy incentive-compatibility.

of initial merit under which a deterministic threshold mechanism is optimal.[14] This mechanism deterministically funds the applicants whose final merit exceeds the mechanism's acceptance threshold; all others are rejected. The acceptance threshold is higher than the grantmaker's preferred threshold. That is, the grantmaker commits to rejecting some applicants who the grantmaker would actually like to fund. This commitment to being tough generates powerful investment incentives.

The condition on the distribution of initial merit entails that the grantmaker's preferred threshold lie in the upper tail of the distribution. In this sense, the grantmaker only wants to approve exceptional applicants and competition is fierce.

The optimal mechanism has an interesting alternative interpretation (Augias and Perez-Richet, 2023, Section 5.3): instead of committing to being tough, the grantmaker allocates optimally based on a *noisy* evaluation of final merit. One way of implementing the optimal mechanism via a noisy evaluation is as follows: if an applicant's final merit fails to pass the acceptance threshold, the evaluation only reveals that the merit failed to pass; else if the final merit passes the threshold, the evaluation perfectly reveals the final merit. The coarse evaluation of failure is important for circumventing the grantmaker's lack of commitment to being tough: since the acceptance threshold is strictly higher than the grantmaker's preferred threshold, a non-committed grantmaker would be tempted to approve applicants near the acceptance threshold, upsetting applicants' investment incentives.

The alternative interpretation gives us a sense of how the grantmaker benefits from delegating the evaluation to referees who are tougher than the grantmaker. The interpretation also speaks to the optimal level of noise: to stimulate productive investments, evaluation should be exact at the top but coarse at the bottom.

Let us briefly elaborate on why it is quite subtle that there is an optimal deterministic mechanism. Understanding the induced investment incentives is key. Given a deterministic threshold mechanism, applicants with initial merit narrowly below the threshold will invest to win a grant. However, applicants with initial merit above the threshold do not invest (they are sure to be funded without effort), and neither do applicants with initial merit far below the threshold (reaching the threshold is

[14] Interestingly, a deterministic threshold mechanism is also optimal (under the condition on the merit distribution) if the grantmaker is concerned with welfare (Augias and Perez-Richet, 2023, Proposition 5). Here, welfare means the final merit of the funded applicants, plus the private payoffs of the funded applicants from winning, minus total investment costs. However, the value of the threshold changes relative to the case where the grantmaker only cares about final merit.

too costly). The grantmaker could try to generate better investment incentives by approving *randomly* for final merit levels near the threshold. If an applicant's initial merit is above the threshold, they are now incentivized to invest so as to escape the lottery. Of the applicants with low initial merit who previously did not invest, some are now incentivized to invest a little to become eligible for the lottery. The downside of approving randomly, of course, is that some applicants that the grantmaker would like to approve will be rejected. Further, some applicants who previously invested to push themselves above the threshold are discouraged from investing as they are now only rewarded randomly. The optimality of deterministic threshold mechanisms thus involves carefully analyzing how these forces balance.

Further reading There is a vast literature on contests studying how competitive forces shape incentives to exert effort. This literature has a lot to say about how prize architecture—are there many small or few large prizes?—shapes investment incentives. The results are nuanced; see, for example, classic work by Moldovanu and Sela (2001) or more recent work by Fang et al. (2020). Morgan et al. (2022) study the impact of evaluation noise on investment incentives, while Kim et al. (2023) study a setting where contestants choose their incentives and how noisily they are evaluated. For further reading, see Fu and Wu (2019) and Vojnović (2015). An important caveat is that the contest literature typically focuses on the aggregate effort across all applicants, including those that go unfunded. This focus potentially limits the applicability to grant allocation since grantmakers may care more about the effort of those applicants who are funded than those who are unfunded, especially if the projects of unfunded researchers are never realized.

4.5 Falsification

Applicants may not only undertake productive investments in their proposals. They may instead spend considerable falsification efforts to make their proposals appear stronger than they are. This nonproductive effort is socially wasteful, hampers the grantmaker's ability to identify good applicants, and hurts the political sustainability of the grant-awarding institution.

We discuss mechanisms that anticipate such falsification incentives, based on work

by Perez-Richet and Skreta (2022). Roughly speaking,[15] the model of Perez-Richet and Skreta (2022) coincides with that of Augias and Perez-Richet (2023) except that the applicant's effort now only has an impact on the type observed by the grantmaker; the effort has no impact on the grantmaker's actual payoffs.

A high-level insight of Perez-Richet and Skreta (2022) is that optimal mechanisms involve *productive falsification*. Even if everyone can craft a brilliant proposal, doing so will be easier for those with better research ideas. Applicants will give in to the temptation of falsifying their proposals, equipping the grantmaker with an (imperfect) instrument for inferring their underlying merit.

Another insight is that, depending on the magnitude of the falsification costs, the optimal mechanism again involves a lottery. A lottery deters excessive falsification by lowering the reward from falsifying to a high level. An added welfare benefit is that fewer resources are devoted to wasteful falsification, which seems in line with common arguments in favor of lotteries (for example, Gross and Bergstrom, 2019).

Further reading Even though the grantmaker may benefit from productive falsification, Perez-Richet and Skreta (2024) note that mechanisms prone to falsification may be politically unsustainable and generate unfair advantages for applicants with higher gaming abilities. Perez-Richet and Skreta (2024) study optimal mechanisms that are immune to falsification attempts. Li and Qiu (2024) study a related setup with multiple applicants and multiple grants. Gross and Bergstrom (2019) study a related setup where the grantmaker lacks commitment power and argue that lotteries may improve welfare. Myers (2024) considers, among other things, a model based on Gross and Bergstrom (2019) where falsification also has positive externalities.

5 Post-award management

While the prospective evaluation of applicants surely plays an important role, grantmakers and funded researchers continue interacting after the grant is conferred. Post-award management and retrospective evaluation expand the toolbox that grantmakers have at their disposal. Goldstein and Kearney (2020) document the extent of post-award management of grants at ARPA-E; see also Azoulay et al. (2019a).

[15] We use Proposition 1 of Perez-Richet and Skreta (2022) to reinterpret their baseline model.

As suggested earlier, a concern in the post-award stage may be that grantmakers cannot monitor how researchers exactly use their funds (the hidden action problem). We next illustrate, using the simple model of Maurer and Scotchmer (2004), how retrospective evaluation helps mitigate this hidden action problem.[16]

5.1 A simple model of post-award management

The model of Maurer and Scotchmer (2004) unfolds over multiple periods. A researcher in each period needs a grant to carry out their project. When awarded a grant, the researcher obtains an immediate private benefit corresponding to, say, career advances or reputation. Working on the project is costly for the researcher but benefits society through knowledge production. The researcher can also choose to shirk by diverting the funds to other activities. No knowledge is produced whenever the research does not work on the project.

Here, we have a classic hidden action problem: the grantmaker wants the researcher to work, but the researcher has an incentive to shirk as the costs are borne privately.

The model also features a hidden information component. Researchers differ in how likely they are to develop fund-worthy research ideas in future periods. This productivity is privately known to the researcher but not to the grantmaker.

The grantmaker can retrospectively evaluate the researcher. Specifically, at the end of each period, the grantmaker learns whether the researcher worked on the project, and can then decide whether to slash future funding.

This policy of retrospective evaluation is effective whenever the prospect of receiving future funding is sufficiently valuable to researchers. However, this is not the case for all researchers. Only those researchers who are sufficiently optimistic about having fund-worthy ideas in the future are willing to incur the costs of working today. For the other researchers, the threat of losing access to funds has little bite since they do not expect to have many fund-worthy ideas in the future.

The optimal grant-allocation policy in Maurer and Scotchmer (2004) has the following structure. The grantmaker assigns a grant to every first-time applicant. In the following periods, the grantmaker evaluates each researcher's past performance and only awards a grant to researchers who delivered results in the past. It follows that researchers with high productivity in generating ideas will work whenever they hold a

[16] Maurer and Scotchmer discuss various instruments. We focus on their discussion of grants.

grant and receive a grant whenever they apply for one. The other researchers will receive a grant only once and not deliver results. Thus, retrospective evaluation and repeated grant allocation jointly allow the grantmaker to channel funds to productive researchers and provide them incentives to use the grant effectively. In Model Box 3, we formalize these ideas.

Model Box 3: Retrospective evaluation and repeated funding

In the following, we set up the model in Maurer and Scotchmer (2004) and show how retrospective evaluation paired with repeated funding can help align researchers' with grantmakers' incentives. Researchers privately know the probability λ of obtaining a fund-worthy idea in each period. If a researcher works on their idea, they incur an effort cost κ. To work on ideas, researchers need grants. A researcher's immediate private value for being funded is v. Researchers discount future payoffs at rate r. Under the rule where researchers are funded if and only if they worked in all past periods, a researcher of type λ has an incentive to work in the first period if

$$v \leq v - \kappa + \frac{\lambda}{r}(v - \kappa). \qquad \text{(No-Shirking)}$$

By shirking (on the left side), the researcher enjoys the private benefit v at zero cost but loses out on all future net benefits. By working, they pay the cost κ; additionally, they enjoy future benefits from funding equal to the per-idea rent $v - \kappa$ times the probability λ of obtaining an idea in any given period, in perpetuity from the next period (and, thus, divided by the discount rate r). The (No-Shirking)-condition can be rearranged to

$$\lambda \geq \frac{r\kappa}{v - \kappa}.$$

This inequality shows that only sufficiently productive researchers (high λ) have the incentives to work hard for continuous funding.

5.2 The pre-award benefit of post-award management

We just saw that post-award management impacts the earlier allocation stage. We now elaborate on this idea using work by Mylovanov and Zapechelnyuk (2017).

A grantmaker allocates one grant among a group of applicants. Each applicant privately knows their own individual merit. After allocating the grant, the grantmaker learns the funded applicant's merit and can impose a limited penalty. Our interpretation is that the grantmaker can infer the applicant's true merit by observing, say, the number of their publications. To penalize the applicant, the grantmaker can debar them from future calls or terminate the grant prematurely. However, penalization is imperfect. The winner may enjoy some career gain from winning the grant, which the grantmaker cannot undo ex-post. Premature termination has bite when researchers rely on the grant to carry out their work or when they are intrinsically motivated about working on their projects. This feature contrasts with the Maurer-Scotchmer model; there, researchers only care about getting funding.

How can the grantmaker use the threat of ex-post penalization to screen the applicants? Mylovanov and Zapechelnyuk (2017, Theorem 1) derive the following optimal mechanism, taking the form of a *binary shortlisting procedure*.[17] First, the grantmaker announces a threshold. Second, applicants are asked to declare whether their merit lies above the threshold (and they will find it in their best interest to do so truthfully). Third, applicants above the threshold are shortlisted, and applicants below the threshold are shortlisted with some probability strictly less than one. Fourth, a winner is selected uniformly at random from the shortlist (if the shortlist is empty, a winner is selected uniformly at random from the full applicant pool). Finally, if the winner is ex-post found to have misreported their quality, the penalty is triggered.

We highlight two qualitative insights from this result. First, the optimal mechanism involves a lottery. To gain an intuition, consider applicants with merit below the threshold. Applicants are more likely to be shortlisted (and hence win the grant) if they claim to pass the threshold. Therefore, if a low-merit applicant would never win the grant by reporting truthfully, they would have an incentive to misreport, even under the threat of ex-post penalization. By introducing the lottery, the grantmaker incentivizes these low-merit types to report truthfully. Thus, the lottery serves as an incentive device for ensuring truthful reports.

The second insight is that ex-post penalties indeed provide incentives ex-ante. Applicants are more likely to be short-listed (and hence win the grant) if they pass the threshold. Thus the optimal mechanism is more likely to allocate the grant to a

[17] Here, we focus on the optimal mechanism in the case with sufficiently many agents. See Theorem 1 in Mylovanov and Zapechelnyuk (2017) for the general solution.

high-merit researcher than a completely random mechanism.

Further reading Li (2020) studies a nearby setting where the grantmaker has to pay a cost to learn the winner's type ex-post.

5.3 Post-award incentives

We next ask how the structure of grants shapes the incentives of funded researchers. As we have seen in Section 5.1, the nature of grants as upfront payments creates a hidden action problem: conditional on being funded, a researcher may have an incentive to use the funds to serve their own interests rather than those of the grantmaker or society. Maurer and Scotchmer (2004) emphasize that grantmakers can re-align interests by monitoring researchers and, possibly, cutting them off from future funding. However, slashing funding might not be a sufficiently powerful instrument. Researchers can apply for grants from different grantmakers or might not require external funding frequently. Further, it is somewhat stringent to assume that grantmakers can perfectly monitor whether researchers work.

In classical hidden action models (Holmström, 1979), incentives are typically provided through performance-dependent payments. Grants do not usually use outcome-dependent rewards, but we can still gain some insights from this literature. In the following, we will first give an overview of relevant insights from the literature assuming that performance-dependent rewards are available to the funder. Then, we will briefly discuss how such performance-dependent rewards can be interpreted in the context of research grants.

Consider a researcher who has been awarded a grant, and there is no uncertainty about the researcher's merit. Once the grant has been allocated, the researcher must decide how to allocate their resources; for example, whether to spend time and effort on the projects in their proposal. While the grantmaker wants the researcher to devote time to those projects, the researcher has private incentives not to; for example, because the researcher prefers working on alternative projects or traveling to conferences. Generally, a grantmaker cannot observe or verify how a researcher allocates their time. Therefore, the grantmaker must use verifiable outputs that depend on the researcher's choices to provide incentives, such as publications related to the grant projects. Due to the uncertain nature of research, publication outcomes

are not a perfect measure of the researcher's choices, but they are likely correlated. Suppose that, by allocating more time to grant-related projects, a researcher can increase the likelihood of having publications at the end of the grant period. Then, the grantmaker can design a reward scheme for the researcher that depends on publication outcomes. Good publication outcomes serve as a signal of the researcher's having allocated their time as the grantmaker desired and are therefore rewarded. If a researcher has worse publication outcomes, the reward should be lower to prevent the researcher from allocating their time differently.

This basic hidden action model has been extended in many directions. In the context of scientific research, we focus on two relevant modeling ingredients. First, grants usually have long time horizons, and researchers repeatedly decide how to allocate their resources. Second, researchers learn about their projects over time—for example, the feasibility of producing any (publishable) results.

For simplicity, suppose the grantmaker wants to encourage a researcher to work on a project for a specified duration. The successful completion is observable (in the form of journal publications, for example). However, whether and when the researcher successfully completes the project is uncertain, even when working on the project as desired. Therefore, the grantmaker cannot infer whether the absence of a publication at each point is due to misfortune or idleness. How can the grantmaker incentivize the researcher to work on the project?

Let us consider the model of Green and Taylor (2016). The grantmaker desires that the researcher successfully completes the funded project and chooses both the duration of the project and how the researcher is rewarded. The researcher then chooses at each point in time whether or not to work on the project. When working on the project, the researcher forgoes the value of using the funds differently. Naturally, the project can only be completed when the researcher works on the project. To capture the uncertainty in research, the researcher may fail to complete the project even when working on it. Green and Taylor (2016) show that the optimal grant design in this case is as follows. The grant provides funding until a (finite) deadline. The reward for completing the project is time-dependent; the earlier the researcher completes the project, the higher the reward. Intuitively, the researcher has to be compensated for two reasons. First, when working on the project, the researcher forgoes the benefits of allocating the funds differently. Second, upon completing the project before the deadline, the researcher forgoes the private enjoyment from the

funds for the remaining time—this is because the researcher could always decide to use the funds differently and claim that the absence of success was due to bad luck.

To capture the idea that a researcher learns about a project while working on it, consider the following model studied by Moroni (2022). It is uncertain whether the project is feasible. If the project is infeasible, it cannot be completed. Thus, the researcher learns about the project's feasibility while working. Specifically, when working without success, the researcher becomes more pessimistic about the project, and hence also more pessimistic about the prospect of winning whatever reward the grantmaker promised for successful completion. The more pessimistic the researcher is, the higher the reward necessary to incentivize effort. Therefore, one might expect the reward to increase over time in the presence of learning. However, this is not the case. Suppose that the researcher would expect a higher reward tomorrow than today. Then, the researcher would rather delay working on the project today (which is unobserved by the grantmaker) and work on the project tomorrow. Because the researcher did not work today, the researcher did not become more pessimistic about the project. Therefore, the researcher must be compensated for earlier successes. The optimal design features a project deadline and a constant reward.[18]

Halac et al. (2016) show how to incorporate hidden information into this model. Suppose the researcher has better information about the project's feasibility than the grantmaker. In this case, the grantmaker can offer a menu of different grant designs so that the researcher reveals their private information via self-selection. Green and Taylor (2016) show how the optimal grant design changes if completing the project requires an intermediate step that is unverifiable to the grantmaker. In this case, the grantmaker optimally incentivizes the researcher by staging the project into two steps with a soft first and a hard second deadline.

Manso (2011) considers a different variant of the dynamic hidden action model focusing on the researcher's choice between two projects. One project is safe while the other project is riskier; in particular, it is unknown how likely it is to succeed on the risky project, and, in expectation, success is less likely than on the safe project. Manso (2011) shows that to encourage work on the riskier project, the grant design must exhibit tolerance for early failure (reflected, for example, in a longer duration)

[18]The constant reward relies on the agent's not discounting the future. If the agent would discount the future, the optimal reward schedule would be declining over time as the incentive to delay the work is slightly mitigated.

and rewards for long-term success. If early successes are rewarded relatively highly, the researcher is incentivized to work on safe, less creative problems rather than novel but risky problems.

On the empirical side, Myers and Tham (2023), using survey data of US professors, provide evidence that a longer grant duration indeed encourages riskier projects, but only for researchers with long-term job security.[19] In response to larger grants researchers shift resources toward existing projects rather than opening new ones. Myers and Tham (2023) also discuss estimates for the rate at which researchers are willing to substitute duration for size.[20]

The results in the preceding discussion of hidden action models rely on the ability of the grantmaker to design and commit to performance-dependent rewards. Whenever such rewards are infeasible or undesirable, the results provide a benchmark on what would be achievable with these additional instruments. It would be interesting to investigate how a grantmaker can design instruments to resemble performance-dependent payments. Maurer and Scotchmer (2004) provide one example of repeatedly assigning grants of fixed value. One could imagine that the grantmaker can condition the access to or the design of future grants on past performance to approximate the optimal performance-dependent rewards. Gross and Bergstrom (2024) interpret the reward in a moral hazard model normatively as how the scientific community *should* evaluate scientific contributions.

Another strand of the literature on dynamic hidden action models focuses on the use of costly inspections to incentivize researchers to work. Ball and Knoepfle (2024) consider a dynamic model in which a researcher decides whether to work on a project. The grantmaker cannot use time-dependent payments in their model but chooses the timing of costly inspections that reveal the researcher's past choices. The researcher's effort increases the probability that a project is completed by working. Completing the project terminates the grant. How can the grantmaker time inspections in such a way that the researcher will work while economizing on the cost of inspections? Ball and Knoepfle (2024) show that the grantmaker optimally inspects the researcher in fixed time intervals.[21]

[19] Here, when we say a grant has a size of $\$x$ and a duration of n years, we mean that the grantee can freely spend up to $\$x$ within n years. After n years, all unused funds expire.

[20] See Tham (2023) and Tham et al. (2024) for the impact of funding delays on career outcomes.

[21] Ball and Knoepfle (2024) also study the case in which working reduces the probability of a *breakdown*, that is, the failure of the project. In this case, random inspections are optimal.

6 Grant supply and the direction of research

The previous sections shed light on *how* the grantmaker should structure payments and design the allocation rule to achieve a particular goal. We now shift our focus to the choice of *what* to fund. How does a grantmaker's funding of a specific topic impact researchers' incentives and the direction of research?

In practice, grantmakers frequently offer grants on specific topics, so-called mission-oriented grants. Examples include *Requests for Applications (RFAs)* at the NIH (Myers, 2020), and the *SWITCHES* program at ARPA-E (Azoulay and Li, 2021). Further, in the context of investigator-initiated proposals, grantmakers can selectively fund proposals that align most closely with their goals, thereby indirectly steering the direction of research. Investigator-initiated proposals perhaps run a higher risk than mission-oriented grants of attracting proposals that do not closely fit the grantmaker's interests. However, investigator-initiated proposals can more easily leverage researchers' knowledge about the prospects of research projects.

Besides the availability of funding and personal intellectual curiosity, researchers' preferences over questions are shaped by feasibility—how difficult is progress given current knowledge?— and the reward for establishing priority. Importantly, the incentives from feasibility and priority may fail to induce an efficient allocation of researchers to questions due to *congestion* and *competition* effects.

When a new researcher joins a field, the rate of discovery in this field increases. *Congestion effects* arise when this increase is decreasing in the number of incumbent researchers. From a social planner's perspective, it may be preferable for the new researcher to reallocate to a field with fewer incumbent researchers. In theoretical work, Hopenhayn and Squintani (2021) develop this idea in a model of corporate R&D. Since firms cannot acquire property rights to open problems (but only to inventions), there is inefficiently high entry into "hot" areas. While theirs is a model of patents and corporate R&D, a similar line of reasoning may apply to academic research since researchers cannot acquire property rights to open questions. Thus, there may be an economic value to communal norms that informally bestow such property rights.

For projects with high ex-ante potential, *competition effects* arise when many researcher simultaneously tackle the problem and attempt to outrace one another to establish priority. Hill and Stein (2024) model these incentives and provide evidence of a negative effect on research quality, finding that projects with "higher ex-ante

potential generate more competition, are completed faster, and are lower quality."[22]

How can grants improve the allocation of researchers to research questions and curtail congestion and competition effects? Let us speculate on some channels. First, grantmakers can create "hot" areas by offering large grants, thereby pulling researchers away from congested areas. Naturally, implementing this would require an idea of what researchers would otherwise work, an issue we return to momentarily. Second, if unfunded researchers cannot investigate certain questions (due to the costs of specialized equipment, for example), grantmakers effectively place an upper bound on the number of researchers working on a question, reducing both competition and congestion effects. Third, grants may act as a commitment device: a researcher with a grant credibly signals that they have the resources to investigate their area.

How effective are grants at guiding the direction of research? Recent empirical work by Myers (2020) quantifies the costs of steering the direction using RFAs at the NIH. The evidence suggests that researchers do not simply follow the money, as it were, but are much more likely to apply for RFAs that are topically similar to their prior research. Overall, Myers (2020) suggests it is costly to incentivize specific topics. In equilibrium, "RFAs must make more funds available to attract the same number of applications as the investigator-initiated mechanism."

The costs and benefits of investigating a question are difficult to quantify since they depend on the current knowledge frontier. Carnehl and Schneider (2023) develop a model for understanding the evolution of scientific knowledge and the researchers' incentives for pursuing questions. The model conceptualizes knowledge as a set of questions to which the answer is known. The answer to any particular question has a spillover effect on nearby questions through improved conjectures about their answers. Having precise information about answers is valuable since it guides decision-making in practical problems. In this model, the interests of researchers and society are misaligned for two reasons. First, researchers do not fully internalize that uncovering

[22]Their model does not assume that establishing priority is a winner-takes-all-contest, but only that being second yields a lower reward than being first; this aligns with empirical evidence from structural biology (Hill and Stein, 2023). The model of Hill and Stein (2024) builds on earlier theoretical work by Bobtcheff et al. (2017), who cast competition between two researchers as a winner-takes-all contest without endogenous entry. Bobtcheff et al. (2017) show that "when breakthroughs become more frequent, researchers are under increasing competitive pressure and have decreasing incentives to wait and let their ideas mature." That is, the adverse incentives to publish low-quality work to establish priority are especially pertinent when technological progress allows the research community to produce insights quickly.

an answer provides a guiding light for future researchers. Second, researchers bear the costs of uncovering answers; these costs are lower for questions similar to ones with known answers. The misaligned interests lead to inefficiently low novelty in research. Hence this model provides a formal framework for investigating the effects of policy tools (such as RFAs) when researchers have a stronger preference than society for questions close to existing knowledge. Interpreting grants as reductions in the cost of research, grants can induce researchers to investigate more novel questions, even when grants are not topic-specific. In this model, however, such grants cannot incentivize "moonshots"—research on extremely novel questions that guide future researchers to choose more efficiently. To encourage researchers to work on moonshot questions, one may have to resort to other tools, such as research prizes or mission-oriented grants.

7 Concluding remarks

Two high-level insights that emerge from our overview of the economics of grant design:

(1) The case for (partial) randomization is mixed. For example, randomization may arise optimally if grantmakers are concerned with evaluation costs or with deterring applicants from inflating their merit. The case for randomization is weaker when accounting for productive investments.

(2) Intermediate reviews and a staged grant design may help realign the incentives of funded researchers with those of the grantmaker, ensuring that researchers direct their resources toward productive areas.

The economics literature has given relatively little attention to grant funding. By collecting insights that are broadly applicable to allocation problems, we hope this piece inspires follow-up work on grant funding as a particular application. In our view, there is scope for both theoretical and empirical work clarifying the interaction of different stages and economic forces in grant funding (Figures 2 and 3), as well as improving our understanding of the relationship between grants and other funding instruments (Figure 1). Let us give two examples.

(1) We discussed separately the role of applicants' productive and unproductive efforts. One could think that efforts are rarely completely unproductive. Indeed, Myers (2024) uses survey data to argue that applicants' fundraising efforts typically do have some positive externalities. It is interesting to analyze in more

detail the application and evaluation costs inherent to grant funding, and to then draw lessons for the design of funding rules.

(2) As already indicated, there is little prior work relating grants to other funding instruments or, more broadly, relating push to pull incentives in science funding. The COVID-19 pandemic spurred greater interest in the question of optimally combining push and pull incentives, perhaps via entirely novel instruments (Ahuja et al., 2021; Athey et al., 2022; D'Souza and Snyder, 2024).

References

Adda, Jérôme and Marco Ottaviani (2023). "Grantmaking, Grading on a Curve, and the Paradox of Relative Evaluation in Nonmarkets". In: *Quarterly Journal of Economics* 139.2, pp. 1255–1319.

Aghion, Philippe, Mathias Dewatripont, and Jeremy C. Stein (2008). "Academic Freedom, Private-Sector Focus, and the Process of Innovation". In: *RAND Journal of Economics* 39.3, pp. 617–635.

Ahuja, Amrita, Susan Athey, Arthur Baker, Eric Budish, Juan Camilo Castillo, Rachel Glennerster, Scott Duke Kominers, Michael Kremer, Jean Lee, Canice Prendergast, Christopher M. Snyder, Alex Tabarrok, Brandon Joel Tan, and Witold Więcek (2021). "Preparing for a Pandemic: Accelerating Vaccine Availability". In: *AEA Papers and Proceedings* 111, pp. 331–35.

Alon, Noga, Felix Fischer, Ariel Procaccia, and Moshe Tennenholtz (2011). "Sum of Us: Strategyproof Selection from the Selectors". In: *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 101–110.

Andrews, Isaiah and Maximilian Kasy (2019). "Identification of and Correction for Publication Bias". In: *American Economic Review* 109.8, pp. 2766–2794.

Arrow, Kenneth J. (1962). "Economic Welfare and the Allocation of Resources for Invention". In: *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton University Press, pp. 609–626.

Athey, Susan, Juan Camilo Castillo, Esha Chaudhuri, Michael Kremer, Alexandre Simoes Gomes, and Christopher M Snyder (2022). "Expanding Capacity for Vaccines Against Covid-19 and Future Pandemics: A Review of Economic Issues". In: *Oxford Review of Economic Policy* 38.4, pp. 742–770.

Augias, Victor and Eduardo Perez-Richet (2023). *Non-Market Allocation Mechanisms: Optimal Design and Investment Incentives*. arXiv: 2303.11805.

Azoulay, Pierre, Erica Fuchs, Anna P. Goldstein, and Michael Kearney (2019a). "Funding Breakthrough Research: Promises and Challenges of the 'ARPA Model'". In: *Innovation Policy and the Economy* 19.1, pp. 69–96.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat (2019b). "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules". In: *Review of Economic Studies* 86.1, pp. 117–152.

Azoulay, Pierre and Danielle Li (2021). "Scientific Grant Funding". In: *Innovation and Public Policy*. Chicago: University of Chicago Press, pp. 117–150.

Azrieli, Yaron (2024). *Temporary Exclusion in Repeated Contests*. arXiv: 2401.06257.

Balietti, Stefano, Robert L. Goldstone, and Dirk Helbing (2016). "Peer Review and Competition in the Art Exhibition Game". In: *Proceedings of the National Academy of Sciences* 113.30, pp. 8414–8419.

Ball, Ian and Jan Knoepfle (2024). *Should the Timing of Inspections be Predictable?* arXiv: 2304.01385.

Ben-Porath, Elchanan, Eddie Dekel, and Barton L. Lipman (2014). "Optimal Allocation with Costly Verification". In: *American Economic Review* 104.12, pp. 3779–3813.

— (2019). "Mechanisms with Evidence: Commitment and Robustness". In: *Econometrica* 87.2, pp. 529–566.

— (2023). "Sequential Mechanisms for Evidence Acquisition".

Bendiscioli, Sandra, Teo Firpo, Albert Bravo-Biosca, Eszter Czibor, Michele Garfinkel, Tom Stafford, James Wilsdon, and Helen Buckley Woods (June 2022). *The Experimental Research Funder's Handbook*.

Biagioli, Mario (2019). *Galileo's Instruments of Credit: Telescopes, Images, Secrecy*. University of Chicago Press.

Bloch, Francis, Bhaskar Dutta, and Marcin Dziubiński (2023). "Selecting a Winner with External Referees". In: *Journal of Economic Theory*, p. 105687.

Bobtcheff, Catherine, Jérôme Bolte, and Thomas Mariotti (2017). "Researcher's Dilemma". In: *Review of Economic Studies* 84.3, pp. 969–1014.

Bryan, Kevin A and Jorge Lemus (2017). "The Direction of Innovation". In: *Journal of Economic Theory* 172, pp. 247–272.

Bryan, Kevin A and Heidi L Williams (2021). "Innovation: Market Failures and Public Policies". In: *Handbook of Industrial Organization*. Vol. 5. 1. Elsevier, pp. 281–388.

Carnehl, Christoph and Johannes Schneider (2023). *A Quest for Knowledge*. arXiv: 2102.13434.

Castillo, Juan Camilo, Amrita Ahuja, Susan Athey, Arthur Baker, Eric Budish, Tasneem Chipty, Rachel Glennerster, Scott Duke Kominers, Michael Kremer, Greg Larson, Jean Lee, Canice Prendergast, Christopher M. Snyder, Alex Tabarrok, Brandon Joel Tan, and Witold Więcek (2021). "Market Design to Accelerate COVID-19 Vaccine Supply". In: *Science* 371.6534, pp. 1107–1109.

Chakravarty, Surajeet and Todd R. Kaplan (2013). "Optimal Allocation Without Transfer Payments". In: *Games and Economic Behavior* 77.1, pp. 1–20.

Condorelli, Daniele (2012). "What Money Can't Buy: Efficient Mechanism Design with Costly Signals". In: *Games and Economic Behavior* 75.2, pp. 613–624.

D'Souza, Arielle and Christopher M. Snyder (2024). "Can Operation Warp Speed Serve as a Model for Accelerating Innovations Beyond Covid Vaccines?" In: *Entrepreneurship and Innovation Policy and the Economy, Volume 4*. University of Chicago Press.

Dasgupta, Partha and Paul A David (1994). "Toward a New Economics of Science". In: *Research Policy* 23.5, pp. 487–521.

De Clippel, Geoffroy, Herve Moulin, and Nicolaus Tideman (2008). "Impartial Division of a Dollar". In: *Journal of Economic Theory* 139.1, pp. 176–191.

Dworczak, Piotr (2023). "Equity-Efficiency Trade-Off in Quasi-Linear Environments".

Epitropou, Markos and Rakesh Vohra (2019). "Optimal On-Line Allocation Rules with Verification". In: *Algorithmic Game Theory*. Ed. by Dimitris Fotakis and Evangelos Markakis. Springer International Publishing, pp. 3–17.

ERC (2023). "Information for Applicants: Starting and Consolidator Grant 2024 Calls".

Fang, Dawei, Thomas Noe, and Philipp Strack (2020). "Turning Up the Heat: The Discouraging Effect of Competition in Contests". In: *Journal of Political Economy* 128.5, pp. 1940–1975.

Fang, Ferric C. and Arturo Casadevall (2016). *Research Funding: The Case for a Modified Lottery*.

Fischer, Felix and Max Klimm (2014). "Optimal Impartial Selection". In: *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, pp. 803–820.

Franzoni, Chiara and Paula Stephan (2023). "Uncertainty and Risk-Taking in Science: Meaning, Measurement and Management in Peer Review of Research Proposals". In: *Research Policy* 52.3, p. 104706.

Fu, Qiang and Zenan Wu (2019). "Contests: Theory and Topics". In: *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

Galilei, Galileo (1609). "Lettera a Leonardo Donato, Doge di Venezia, 26 Agosto 1609". In: *Le Opere di Galileo Galilei, Vol. X*. Ed. by Antonio Favara. Barbera, 1890 - 1909, pp. 250–251.

— (1610). "Lettera a Belisario Vinta, 7 Maggio 1610". In: *Discoveries and Opinions of Galileo*. Doubleday Anchor Books. Trans. by Drake Stillman as *Letter to Belisario Vinta, 7 May 1610* (1957).

Gallini, Nancy and Suzanne Scotchmer (2002). "Intellectual Property: When Is It the Best Incentive System?" In: *Innovation Policy and the Economy* 2, pp. 51–77.

Goldstein, Anna P. and Michael Kearney (2020). "Know When to Fold 'Em: An Empirical Description of Risk Management in Public Research Funding". In: *Research Policy* 49.1, p. 103873.

Green, Brett and Curtis R Taylor (2016). "Breakthroughs, Deadlines, and Self-Reported Progress: Contracting for Multistage Projects". In: *American Economic Review* 106.12, pp. 3660–3699.

Gross, Kevin and Carl T. Bergstrom (2019). "Contest Models Highlight Inherent Inefficiencies of Scientific Funding Competitions". In: *PloS Biology* 17.1, pp. 1–15.

— (2024). *Rationalizing risk aversion in science*. arXiv: 2306.13816.

Halac, Marina, Navin Kartik, and Qingmin Liu (2016). "Optimal Contracts for Experimentation". In: *Review of Economic Studies* 83.3, pp. 1040–1091.

Hartline, Jason D. and Tim Roughgarden (2008). "Optimal Mechanism Design and Money Burning". In: *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, pp. 75–84.

Heyard, Rachel, Manuela Ott, Georgia Salanti, and Matthias Egger (2022). "Rethinking the Funding Line at the Swiss National Science Foundation: Bayesian Ranking and Lottery". In: *Statistics and Public Policy* 9.1, pp. 110–121.

Hill, Ryan and Carolyn Stein (2023). "Scooped! Estimating Rewards for Priority in Science".

— (2024). "Race to the Bottom: Competition and Quality in Science".

Hippel, Ted von and Courtney von Hippel (2015). "To Apply or Not to Apply: A Survey Analysis of Grant Writing Costs and Benefits". In: *PloS One* 10.3, e0118494.

Holmström, Bengt (1979). "Moral Hazard and Observability". In: *Bell Journal of Economics*, pp. 74–91.

Holzman, Ron and Hervé Moulin (2013). "Impartial Nominations for a Prize". In: *Econometrica* 81.1, pp. 173–196.

Hopenhayn, Hugo and Francesco Squintani (2021). "On the Direction of Innovation". In: *Journal of Political Economy* 129.7, pp. 1991–2022.

Hussam, Reshmaan, Natalia Rigol, and Benjamin N Roth (2022). "Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design in the Field". In: *American Economic Review* 112.3, pp. 861–898.

Jones, Benjamin F. (2021). "Science and Innovation: The Under-Fueled Engine of Prosperity". In: *Rebuilding the Post-Pandemic Economy*. Ed. by Melissa S. Kearney and Amy Ganz. Washing D.C.: Aspen Institute Press.

Kattwinkel, Deniz (2020). "Allocation with Correlated Information: Too Good to Be True". In: *Proceedings of the 21st ACM Conference on Economics and Computation*, pp. 109–110.

Kattwinkel, Deniz and Jan Knoepfle (2023). "Costless Information and Costly Verification: A Case for Transparency". In: *Journal of Political Economy* 131.2, pp. 504–548.

Kattwinkel, Deniz, Axel Niemeyer, Justus Preusser, and Alexander Winter (2022). "Mechanisms Without Transfers for Fully Biased Agents". In: *Proceedings of the 23rd ACM Conference on Economics and Computation*, pp. 332–333.

Khalfan, Nawaaz (2023). "Optimal Allocation with Noisy Inspection".

Kim, Kyungmin, R. Vijay Krishna, and Dmitry Ryvkin (2023). "Choosing Your Own Luck: Strategic Risk Taking and Effort in Contests".

Kurokawa, David, Omer Lev, Jamie Morgenstern, and Ariel D Procaccia (2015). "Impartial Peer Review". In: *IJCAI*. Vol. 15, pp. 582–588.

Lehmann, Erich Leo (1988). "Comparing Location Experiments". In: *Annals of Statistics* 16.2, pp. 521–533.

Li, Danielle, Pierre Azoulay, and Bhaven N Sampat (2017). "The Applied Value of Public Investments in Biomedical Research". In: *Science* 356.6333, pp. 78–81.

Li, Yingkai and Xiaoyun Qiu (2024). *Screening Signal-Manipulating Agents via Contests*. arXiv: 2302.09168.

Li, Yunan (2020). "Mechanism Design with Costly Verification and Limited Punishments". In: *Journal of Economic Theory* 186, p. 105000.

Li, Zihao and Jonathan Libgober (2023). "The Dynamics of Verification When Searching for Quality".

Lotka, Alfred J (1926). "The Frequency Distribution of Scientific Productivity". In: *Journal of the Washington Academy of Sciences* 16.12, pp. 317–323.

Mackenzie, Andrew (2015). "Symmetry and Impartial Lotteries". In: *Games and Economic Behavior* 94, pp. 15–28.

Manso, Gustavo (2011). "Motivating Innovation". In: *Journal of Finance* 66.5, pp. 1823–1860.

Matcham, William and Mark Schankerman (2023). "Screening Property Rights for Innovation".

Maurer, Stephen M. and Suzanne Scotchmer (2004). "Procuring Knowledge". In: *Intellectual Property and Entrepreneurship*. Emerald Group Publishing Limited, pp. 1–31.

Merrifield, Michael R and Donald G Saari (2009). "Telescope Time Without Tears: A Distributed Approach to Peer Review". In: *Astronomy & Geophysics* 50.4, pp. 4–16.

Merton, Robert K. (1957). "Priorities in Scientific Discovery: A Chapter in the Sociology of Science". In: *American Sociological Review* 22.6, pp. 635–659.

Mervis, Jeffrey (2014). "A Radical Change in Peer Review". In: *Science* 345.6194, pp. 248–249.

Moldovanu, Benny and Aner Sela (2001). "The Optimal Allocation of Prizes in Contests". In: *American Economic Review* 91.3, pp. 542–558.

Morgan, John, Justin Tumlinson, and Felix Várdy (2022). "The Limits of Meritocracy". In: *Journal of Economic Theory* 201, p. 105414.

Moroni, Sofia (2022). "Experimentation in Organizations". In: *Theoretical Economics* 17.3, pp. 1403–1450.

Myers, Kyle (2020). "The Elasticity of Science". In: *American Economic Journal: Applied Economics* 12.4, pp. 103–134.

Myers, Kyle and Wei Yang Tham (2023). *Money, Time, and Grant Design*. arXiv: 2312.06479.

Myers, Kyle R. (2024). *Some Tradeoffs of Competition in Grant Contests*. arXiv: 2207.02379.

Mylovanov, Tymofiy and Andriy Zapechelnyuk (2017). "Optimal Allocation with Ex Post Verification and Limited Penalties". In: *American Economic Review* 107.9, pp. 2666–2694.

Niemeyer, Axel and Justus Preusser (2023). "Simple Allocation with Correlated Types".

Olckers, Matthew and Toby Walsh (2023). *Manipulation and Peer Mechanisms: A Survey*. arXiv: 2210.01984.

Olszewski, Wojciech (2020). "A Theory of Citations". In: *Research in Economics* 74.3, pp. 193–212.

Pasachoff, Jay M (2015). "Simon Marius's Mundus Iovialis: 400th Anniversary in Galileo's Shadow". In: *Journal for the History of Astronomy* 46.2, pp. 218–234.

Pereyra, Juan S. and Francisco Silva (2023). "Optimal Assignment Mechanisms with Imperfect Verification". In: *Theoretical Economics* 18.2, pp. 793–836.

Perez-Richet, Eduardo and Vasiliki Skreta (2022). "Test Design Under Falsification". In: *Econometrica* 90.3, pp. 1109–1142.

— (2024). "Fraud-Proof Non-Market Allocation Mechanisms".

Price, W. Nicholson (2019). "Grants". In: *Berkeley Technology Law Journal* 34.1, pp. 1–66.

Shah, Nihar B., Behzad Tabibian, Krikamol Muandet, Isabelle Guyon, and Ulrike Von Luxburg (2018). "Design and Analysis of the NIPS 2016 Review Process". In: *Journal of Machine Learning Research* 19.49, pp. 1–34.

Stephan, Paula (2012). *How Economics Shapes Science*. Harvard University Press.

Stern, Scott (2004). "Do Scientists Pay to Be Scientists?" In: *Management Science* 50.6, pp. 835–853.

Tham, Wei Yang (2023). "Science, Interrupted: Funding Delays Reduce Research Activity but Having More Grants Helps". In: *Plos One* 18.4, pp. 1–40.

Tham, Wei Yang, Joseph Staudt, Elisabeth Ruth Perlman, and Stephanie D. Cheng (2024). *Scientific Talent Leaks Out of Funding Gaps*. arXiv: 2402.07235.

Vojnović, Milan (2015). *Contest Theory: Incentive Mechanisms and Ranking Methods*. Cambridge University Press.

Wright, Brian D (1983). "The Economics of Invention Incentives: Patents, Prizes, and Research Contracts". In: *American Economic Review* 73.4, pp. 691–707.

Yang, Yang, Tanya Y. Tian, Teresa K. Woodruff, Benjamin F. Jones, and Brian Uzzi (2022). "Gender-Diverse Teams Produce More Novel and Higher-Impact Scientific Ideas". In: *Proceedings of the National Academy of Sciences* 119.36, e2200841119.

Yoon, Kiho (2011). "Optimal Mechanism Design When Both Allocative Inefficiency and Expenditure Inefficiency Matter". In: *Journal of Mathematical Economics* 47.6, pp. 670–676.