GRANTMAKING, GRADING ON A CURVE, AND THE PARADOX OF RELATIVE EVALUATION IN NONMARKETS*

JÉRÔME ADDA AND MARCO OTTAVIANI

The article develops a model of nonmarket allocation of resources such as the awarding of grants to meritorious projects, honors to outstanding students, or journal slots to quality publications. On the supply side, the available budget of grants is awarded to applicants who are evaluated most favorably according to the noisy information available to reviewers. On the demand side, stronger candidates are more likely to obtain grants and thus self-select into applying, given that applications are costly. We establish that if evaluation is perfect, grading on a curve inefficiently discourages even the very best candidates from applying. More generally, when the budget is insufficient to award grants to all applicants, the equilibrium unravels if information is symmetric enough—the paradox of relative evaluation. Leveraging a technique based on the quantile function pioneered by Lehmann, we characterize a broad set of nonmarket allocation rules under which an increase in evaluation noise in a field (or course) raises equilibrium applications in that field, and reduces applications in all other fields. We empirically confirm these comparative statics by exploiting a change in the rule for apportioning the total budget to applications in different fields at the European Research Council, showing that a 1 standard deviation increase in own evaluation noise leads to a 0.4 standard deviation increase in the number of applications and budget share. Moreover, we derive insights for the design of evaluation institutions, particularly regarding the endogenous choice of noise by fields or courses and the optimal aggregation of fields into panels. *JEL codes:* D83, H81, I23.

* We thank without implication Philippe Aghion, Manuel Arellano, Ricardo Alonso, Christian Bjerke, Estelle Cantillon, Mikhail Drugov, Christian Dustmann, Albin Erlanson, Irwin Feller, Alfonso Gambardella, Nicola Gennaioli, Ian Jewitt, Charlie Johnson, Nenad Kos, Danielle Li, Massimo Marinacci, Andreu Mas-Colell, Meg Meyer, Andrea Prat, Ron Siegel, Timothy Simcoe, William Thomson, Giovanni Ursino, Reinhilde Veugelers, Huseyin Yildirim, John Walsh, Glen Weyl, and Richard Zeckhauser for helpful discussion; Kristin Oxley and Eystein Strømmen for data assistance at the Research Council of Norway; Julien Adda, James Atkins, Mohamed Badaoui, Francesco Bilotta, Aldo Lucia, Marta Mojoli, and Massimiliano Pozzi for outstanding research assistance; and Julien Manili, Anna Merotto, Federico Pessina, Eugenio Piga, Maik Sälzer, Nicolas Sourisseau, and Biao Yang for proofreading. We gratefully acknowledge research support by the NBER-Sloan Foundation Science of Science Funding program, the European Patent Office (EPO) Academic Research Programme, and the European Research Council through Grant 101055295 InfoEcoScience (Information Economics for Science).

The Quarterly Journal of Economics (2024), 1255–1319. https://doi.org/10.1093/qje/qjad046.
Advance Access publication on September 18, 2023.

"The just, then, is a species of the proportionate…proportion is equality of ratios

$$\frac{B_i}{B_j} = \frac{a_i}{a_j} \text{ and, therefore, alternando } \frac{B_i}{a_i} = \frac{B_j}{a_j}$$

…all men agree that what is just in distribution must be according to merit in some sense, though they do not all specify the same sort of merit, but democrats identify it with the status of freeman, supporters of oligarchy with wealth (or with noble birth), and supporters of aristocracy with excellence."

—Aristotle, *Nicomachean Ethics*, Book V, chapter 3

I. Introduction

Over the sweep of history, artists and scientists have long relied on wealthy patrons and public support to finance their inventions and discoveries. In 1610 Galileo Galilei wrote to his former pupil Cosimo de' Medici, the Grand Duke of Tuscany, subtly asking for financial support to explore the sky with his new powerful telescope. To lure the patron, Galileo named Jupiter's moons the Medician stars and promised "many discoveries and such as perhaps no other prince can match." Cosimo was duly impressed and granted Galileo a full teaching buyout at the University of Pisa.[1]

A more systematic process for funding talented scholars emerged in embryonic form in the first half of the nineteenth century when science academies in France and England started offering encouragements and grants to support worthy projects by their members.[2] To ensure the best use of funds, learned societies began formalizing the application cycle and the review process to select grant recipients. Similar selection procedures had been in place for centuries at university colleges for assigning scholarships to promising students from families with limited means.[3]

With its roots steeped in patronage, grantmaking evolved in the modern era to become an effective method for identifying prospects worthy of funding. As Carnegie, Rockefeller, Russell Sage, and other industrial tycoons turned philanthropists at the

1. The quote from Galileo is reported in Westfall (1985, 22). For more on Galileo's patronage, see Biagioli (1990) and references therein.

2. See MacLeod (1971) and Crosland and Gálvez (1989).

3. Rashdall (1895, 200–204) describes the examination procedures for selecting applicants at the first university college, the College of Spain, founded at the University of Bologna from a bequest in 1367 and still active today.

beginning of the twentieth century, the private foundations they endowed to "promote the well-being of humankind" were inundated with requests for donations. Leveraging their business experience, trustees of these large foundations refined grantmaking as a systematic approach to "wholesale" giving. Modern philanthropic foundations select which applications to fund with the assistance of specialized evaluation panels and delegate to grantees the "retail" implementation of the charitable work.[4]

As World War II drew to a close, John Maynard Keynes (1945) stewarded the adoption of the grantmaking model with the creation of the Arts Council of Britain by the UK government to "stimulate, comfort and support" independent artistic initiatives in drama, music, and painting.[5] At around the same time in the United States, Bush (1945), building on his success as director of the wartime Office of Scientific Research and Development, forcefully argued in favor of federal support of the best, curiosity-driven "basic research in the colleges, universities, and research institutes" for a wide range of sciences. In 1946 the National Institutes of Health (NIH) greatly expanded its extramural grants program to cover all areas of biomedical research, and in 1950, the National Science Foundation (NSF) was established to fund basic research across a broad range of scientific disciplines.

As grantmaking grew exponentially in the postwar period, funding organizations developed structured procedures for soliciting and evaluating grant applications.[6] Even though expert peer reviewers are capable of evaluating projects in their specialized area, they tend to advocate for increased funding in their field of expertise, to the detriment of other fields. The need to reconcile these conflicting requests makes the allocation of budget across diverse fields particularly thorny. Universities face similar challenges when making decisions about the allocation of resources and positions across departments.

4. See Zunz (2012) and Leat (2016).

5. The U.S. Congress chartered the National Endowment for the Arts in 1965. The grantmaking model for supporting the arts has since been adopted by governments throughout the world, both at the local and national levels; see Upchurch (2016).

6. Nowadays, U.S. federal institutions such as the NIH or the NSF fund research across various fields, amounting to approximately $172 billion a year. The Horizon Europe program has a funding budget of €95.5 billion for the period from 2021 to 2027.

A number of major funders, like the NIH, the main Canadian funding agencies, and the European Research Council (ERC), adopt a "bottom-up" approach based on an apportionment formula that allocates the total available budget to each field depending on the applications received in all fields.[7] These research funding organizations allocate their budget B in proportion to the number of applications a_1, a_2, ..., a_N received in each field $i = 1$, 2, ..., N, resulting in budget

$$B_i = \frac{a_i}{\sum_{j=1}^{N} a_j} B \tag{PA}$$

for field i.[8] Under proportional apportionment (PA), fields vie against one another for funding based on the quantity of applications they attract. It is therefore important to understand what drives applications across different fields and how sensitive the funding of a given field is to what happens in other fields. Answering these questions is crucial for improving the design of the funding process.

A recent change in the budget allocation rule at the ERC, the largest research funding organization in the EU, had significant effects on applications and the final distribution of funding. The ERC organizes its panels into three disciplinary domains: Life Sciences (LS), Physical Sciences and Engineering (PE), and Social Sciences and Humanities (SH). Before 2014, ERC funds were allocated according to PA in proportion to applications received by panels belonging to the same disciplinary domain. Starting in 2015, the PA formula was applied across the board so that each panel's budget became proportional to applications received by that panel relative to applications received by all panels regardless of the domain, rather than relative only to applications received by the panels in the same domain as before. As shown in Figure I, the reform was followed by a substantial change in relative applications and budget shares across domains (as well as panels within the different domains), with a 60% increase in the funding for SH panels and a 20% decrease for LS panels.

7. See Azoulay et al. (2019) for a description of the organization of NIH funding and for a quantification of the effect of NIH research support on innovation through patenting activity.

8. See, e.g., European Commission (2007). In a more general version of the formula, a_i represents the sum of the possibly variable amounts requested by all applicants in field i. However, most applicants tend to apply for the maximum amount allowed.

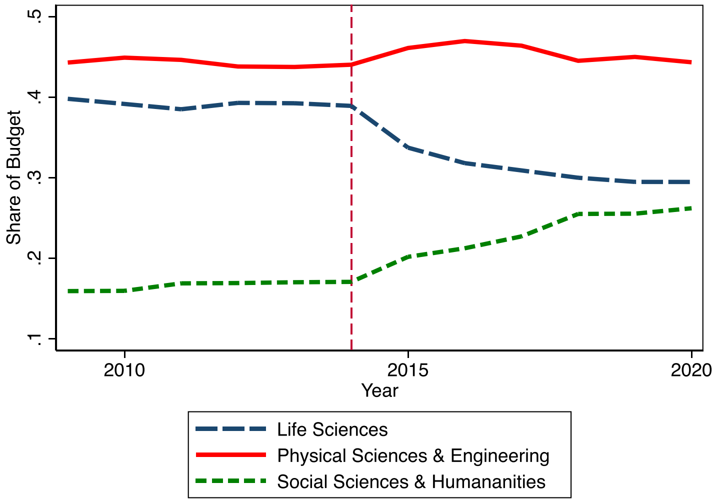

FIGURE I

Budget Shares in ERC Funding by Disciplinary Domain

This figure shows the evolution of the budget shares for the three disciplinary domains at the ERC. The vertical dashed line indicates the last year before the 2014 reform. *Source:* ERC data.

To understand how budget allocation rules and evaluation noise affect the incentives to apply across fields, this article formulates a foundational model of nonmarket allocation of resources. Activities are characterized by their ex ante uncertain merit type, which captures the social value from financing the activity. Proponents of activities can come forth by applying at a cost. The review panel in each field then evaluates and ranks applications based on noisy information to select the most worthy activities. Evaluation is noisy because it involves an important component of expert judgment with subjective evaluation.

As we argue, a key element determining the incentives to apply in different fields is the relative level of evaluation noise, which tends to vary systematically across fields. In some fields, researchers are likely to agree on the quality and novelty of projects, other fields that lack a shared paradigm are characterized by more disagreement. The article derives a general set of budget allocation rules (generalizing PA) under which an increase in evaluation noise in a field increases the applications for grants in that field and its funding at the expense of other fields.

According to our headline result, in fields with little evaluation noise, researchers who are not at the cutting edge refrain from applying because they stand a low chance of being funded, thus reducing available funds in those fields. In contrast, noisy fields receive more applications and therefore obtain relatively more funds.

Our framework encompasses essential characteristics found in a diverse array of nonmarket resource allocation problems. These include the admission process for students across various courses or university degree programs, the selection of articles for publication in journals, the allocation of funds for business projects in conglomerates, and the determination of individuals or businesses to support through government grant-in-aid programs. The following section provides a roadmap of the article, outlining the key points of our analysis primarily in the context of science funding, with brief references to other applications along the way. Section IX casts our contribution in the literature. Section X concludes.

II. Roadmap and Main Insights

II.A. Grantmaking in a Single Field

Section III sets the stage by analyzing the baseline specification with a single field populated by a continuum of candidates parameterized by their merit type. Submitting an application is costly but allows the applicant to gain a private benefit when obtaining a grant. The evaluator appraises the merit of each application received based on a noisy signal—allowing for imperfect information is essential to justify the fact that many applicants do not succeed in obtaining grants. We model the information (or noise) content of the signal through a quantile function approach pioneered by Lehmann (1988), which we leverage for equilibrium comparative statics.[9]

Given the limited budget available for distribution in the field, grants are supplied to the applications that receive sufficiently favorable evaluations. In turn, the evaluation on the supply side induces candidates to apply only when they perceive a chance of success that is sufficiently high to compensate for

9. While the approach was developed for the purpose of welfare comparisons, our analysis showcases the advantage of using Lehmann (1988) for equilibrium comparative statics, relative to Blackwell's (1951) common notion of information.

the application cost. Because higher-merit applicants receive more favorable evaluations, on the demand side, candidates with a merit type above a threshold self-select into applying. We establish the following key comparative statics result: as evaluation becomes noisier, the probability of winning a grant becomes less responsive to the applicant's type, thus increasing the equilibrium amount of applications for a given budget.

II.B. Percentiling and Grading on a Curve

When the allocation across fields is based on raw scores, specialized panels in each field have an incentive to inflate scores to attract greater resources to their respective field. To counteract the resulting grade inflation across panels, from 1988, the NIH started percentiling scores in each panel (known as study section at the NIH) and introduced the payline system.[10] In each panel, grants are assigned to projects that obtain percentiled scores above a level, known as payline, that is equalized across panels. Note that the payline system is equivalent to proportional allocation, given that PA implies that the success rate in field i, defined as the fraction of successful applications in the field i

$$p_i = \frac{B_i}{a_i} = \frac{B}{\sum_{j=1}^{N} a_j}, \tag{1}$$

is automatically equalized across all fields, $p_i = p$. Expert evaluators in each panel are then asked to select the most fund-worthy applications so as to exhaust $100 \times p$ percent of the budget requested by the applications in the field.

Similarly, teachers have incentives to give high grades to students to increase enrollment in their classes to the benefit of their department; see Johnson (2013).[11] With grading on a curve, a constant fraction p of student enrollment a in a class (or

10. See Mandel's (1996, 182–188) historical account. "A percentile ranks your application relative to the other applications reviewed by your study section. …Percentiling counters a phenomenon called "score creep" where study sections give applications increasingly better scores. As a result, scores cluster in the exceptional range, making it impossible to discriminate among applications. Each study section can apply the NIH review criteria differently, scoring either more harshly or more favorably. Percentiling counters these trends by ranking applications relative to others scored by the same study section." https://www.niaid.nih.gov/grants-contracts/understand-paylines-percentiles.

11. Relative grading can also be induced by regulation. For example, according to Texas's Top 10% Rule, students who graduate in the top 10% of their high

degree program) can be awarded honors, so the budget of awards pa is proportional to enrollment. The case with constant payline directly captures grading on a curve for a course that awards a given fraction of distinction grades or honors to enrolled students.

As shown in Section III.B, when the evaluator's signal is additive in the applicant's type, the constant-payline equilibrium is unique and stable if the type distribution has an increasing hazard rate. Multiple equilibria arise only when the type distribution has a segment with decreasing hazard rate, so that the marginal type that is added to the applicants' pool as applications increase becomes closer to the average type of the inframarginal applicants.

II.C. Paradox of Relative Evaluation

As the evaluation signal in a field becomes less noisy, applications in that field in all stable equilibria unambiguously decrease—and decrease more under constant payline than with a fixed budget. Consider the limit case where the grantmaker can perfectly evaluate applicants' merit types without noise. Under a constant payline, only a fraction $p < 1$ of applicants win. With perfect information, candidates know their ranking. Candidates not in the top $100 \times p$ percent of the applicant's pool anticipate that they have no chance of succeeding and thus hold off to save the application cost. Iterating the logic, when the evaluation is perfect, the equilibrium always unravels: no candidate applies in the only outcome compatible with equilibrium. Reversing the logic leading to a market breakdown in Akerlof (1970), here, good types, when perceived as such, make competition for scarce grants tougher and thus drive out bad types. But as applications decrease, the pool of grants is proportionally reduced so that top types dig their own grave. Remarkably, with relative evaluation, symmetric information leads to breakdown. While trading breaks down in classic market settings when information is asymmetric, in our nonmarket environment, information asymmetry is needed to avoid breakdown. This is the paradox of relative evaluation.

More subtly, the equilibrium unravels when the evaluator's signal is sufficiently informative, provided that the hazard rate of the type distribution is bounded, even when the hazard rate is increasing (e.g., with logistic types). When the type distribution

school class are guaranteed automatic admission to state-funded universities. See Cullen, Long, and Reback (2013) for an empirical analysis.

has a vanishing hazard rate, as in the Weibull distribution with a tail thicker than exponential, there is a stable equilibrium with unraveling for any level of noise—and the unraveling equilibrium is unique when information is sufficiently symmetric.

II.D. Partial Equilibrium

When a field is sizable, under PA, an increase in applications in a field reduces the success rate, even when applications in the other fields are held constant. Our general analysis of the partial equilibrium characterizes the allocation resulting in a nonnegligible field.[12] We show that when the payline decreases in applications, uniqueness and comparative statics are preserved when the type distribution has an increasing hazard rate. Uniqueness is lost with a decreasing hazard rate, but all the stable equilibria retain our comparative statics—applications increase in noise.

II.E. General Equilibrium across Fields

Building on the partial equilibrium analysis, Section IV turns to grantmaking across fields where applicants in each field are possibly characterized by different parameters: application cost and award benefit, type and signal distributions, and noise in the evaluator signal. The general equilibrium takes into account the supply-side interdependence through the budget allocation rule. We derive conditions for subproportional budget allocation rules (encompassing fixed budget and PA as special cases) under which equilibrium applications in a field increase when the evaluation in the same field becomes noisier and decrease when the evaluation in other fields becomes noisier.

II.F. Empirical Validation

Leveraging the 2014 reform of the ERC budget apportionment rule, Section V empirically tests the key comparative statics prediction about the effect of noise on applications. This change in the apportionment rule allows us to identify the effect of evaluation noise on the number of applications in each field, relying on a difference-in-differences design. To that effect, we provide

12. This case is analogous to a partial equilibrium analysis in an international trade model, where the country is large enough to affect the terms of trade. The constant payline case corresponds to the partial equilibrium for a small country. In our context, a field does not affect the payline when it has negligible applications relative to the other fields.

novel evidence on evaluation noise across fields using unique data on reviewer grades of grant applications at the Research Council of Norway (RCN). We find stark differences in evaluation noise across fields, with social sciences and applied sciences displaying more noise. We show that the relative differences in evaluation noise across fields significantly predict changes in the number of applicants following the 2014 reform, with sizable and policy-relevant effects. A 1 standard deviation increase in evaluation noise in a given field leads to an increase in the budget for that field of about 0.4 standard deviations.

II.G. Endogenous Evaluation Noise: Game among Fields

Section VI endogenizes the level of evaluation noise in different fields by analyzing a game played by field representatives defending the professional interests of their field. Each field, acting as a collective through this representative, has some capacity to introduce noise in the evaluation of their field, for example, by affecting the quality of the panel members or introducing some randomization in the signal obtained by the panelists.[13]

If field representatives care about the quality of the research financed in their respective fields, they face a trade-off. Increasing noise increases applications but reduces the average quality of the projects selected. In the resulting Nash equilibrium of the game, fields add noise provided that the initial noise is not too high. When the initial noise is already high in at least one of the fields, the noise across fields in the Nash equilibrium remains asymmetric as in our baseline analysis, showing robustness. When the initial noise is asymmetric and relatively low, in equilibrium noise increases in all fields but is also equalized, thus neutralizing the initial asymmetry. When, in addition to being relatively low, the initial noise is sufficiently asymmetric across fields, the final Nash equilibrium allocation results in higher social welfare than is achieved in the (highly inefficient) initial asymmetric allocation.

II.H. Sorting across Fields/Courses

Section VII extends the analysis to incorporate the demand-side interdependence generated by the ability of candidates to select courses/fields depending on their chance of obtaining a high

13. In the case of the ERC, the representative could be the chair of the field panel or the member of the scientific council more closely associated with the field.

grade. Grades have a discernible effect on the future of students (Murphy and Weinhardt 2020). Given that students tend to select courses where they expect to obtain better grades, instructors have incentives to grade generously (Achen and Courant 2009). To curb grade inflation, universities respond by limiting the fraction of students who can obtain top grades and honors (Johnson 2013).

In the baseline model, candidates choose whether to apply/enroll; in this extension, they choose one of two courses, for example, physics or literature, by comparing their chance of ranking among the top $100\times p$ candidates. In the spirit of Roy's (1951) model of occupational sorting, suppose that candidates have a two-dimensional type corresponding to their mathematical and verbal skills. Holding fixed acceptance/merit standards, a field attracts more talented candidates when its evaluation becomes less noisy—intuitively, less talented candidates prefer to hide in the noisier field. As we argue, in equilibrium applications increase in the field's noise and decrease in the noise in the other field. The effect is stronger with grading on a curve than under a fixed budget.

II.I. Design of Funding Rules

Section VIII turns to organizational design questions. Section VIII.A compares the baseline model's general equilibrium allocation with the evaluator's optimal allocation. The optimal amount of applications in a field increases in the evaluation noise in another field, contrary to what happens in the equilibrium induced by a subproportional allocation rule. Starting from the symmetric allocation resulting when fields have symmetric parameters, general equilibrium applications in a field increase excessively in noise relative to the socially optimal allocation. Evaluator welfare can then be improved by decreasing proportionality in fields characterized by less noisy evaluation.

II.J. Pooling Fields and Benchmarking

Section VIII.B considers the effect of pooling a noisier field with a more consensual field into a single panel. Supposing applicants are still evaluated in the same way, what matters for funding once fields are pooled is the candidate's position in the mixture distribution of scores in the two fields. Now candidates evaluated with more (or less) noise are less (or more) likely to be at the top of the distribution. Intuitively, more accurate information increases the scores' dispersion, which matches the underlying type distribution more closely. This way, the more accurate field gains the lion's share of grants in the pooled panel, at the expense of the noisier field. Pooling fields with heterogeneous noise thus dampens the perverse effect of meritocracy on relative evaluation.

III. Grantmaking in a Single Field

This section formulates our baseline grantmaking model in a single field. The field is populated by a continuum of candidates parametrized by their merit type θ, corresponding to the value created if the project is financed. Candidates know their merit, which follows the cumulative distribution G in the population, with size normalized to one. For convenience assume that G admits a continuously differentiable and strictly positive density g on a connected support $[\underline{\theta}, \overline{\theta}]$, possibly unbounded on either side.[14]

To be considered for a grant award, candidates must apply at a cost of c, the opportunity cost of the time spent preparing the application and describing the work.[15] Applicants awarded grants obtain a private benefit v in terms of career advancement and kudos.[16]

An evaluator (review panel) allocates a budget of grants B to applicants on the basis of a noisy signal x about the merit type θ of each applicant. The signal is distributed according to

$$F_\sigma(x \mid \theta), \tag{2}$$

14. We have $G^{-1}(0) = \underline{\theta} = -\infty$ if the support is unbounded below and $G^{-1}(1) = \overline{\theta} = \infty$ if the support is unbounded above.

15. Application costs can be sizable. According to survey evidence by von Hippel and von Hippel (2015) on astronomers and social and personality psychologists who submitted applications for basic research grants to NASA, the NIH, and the NSF, principal investigators spent on average 116 hours preparing applications. This represents a major increase to the early day of science funding. For comparison, in 1921, the prominent German biochemist Otto Warburg submitted to the Notgemeinschaft der Deutschen Wissenschaft (Emergency Association of German Science, the forerunner of the Deutsche Forschungsgemeinschaft) a funding application with a single sentence: "I require 10,000 marks"; see Koppenol, Bounds, and Dang (2011).

16. The model can also easily accommodate the addition of an embarrassment or psychological cost d borne by the candidate when the application is turned down. The cost-benefit ratio $\frac{c}{v}$, which determines demand incentives, is then replaced by $\frac{c+d}{v+d}$.

with continuously differentiable and strictly positive density f_σ and connected support, possibly unbounded on either side. We assume that the signal satisfies the monotone likelihood ratio (MLR) property

$$\text{(MLR)} \qquad \frac{f_\sigma(x\,|\,\theta')}{f_\sigma(x\,|\,\theta)} \text{ increases in } x \text{ for any } \theta' > \theta,$$

so that a higher signal indicates higher merit. A key role in our analysis is played by the parameter σ, which measures the noise in the signal in the following nonparametric way:

$$F_\sigma\left(F_\sigma^{-1}(q\,|\,\theta)\,|\,\theta'\right) \text{ increases in } \sigma \text{ for any } \theta' > \theta$$
$$\text{and any percentile } q \in [0,1]. \tag{3}$$

Lemma 1 in Online Appendix A verifies that an increase in noise according to criterion (3) corresponds to a reduction in information in the sense of Lehmann (1988).[17]

For algebraic tractability, we often illustrate our results for the special case with additive noise where the signal has a location-scale structure, $x = \theta + \sigma\varepsilon$, with noise distribution $F(\varepsilon) = F(\frac{x-\theta}{\sigma})$ and support $[\underline{\varepsilon}, \overline{\varepsilon}]$, possibly unbounded on either side. The signal perfectly reveals the merit when $\sigma = 0$ and becomes completely uninformative as $\sigma \to \infty$.[18]

Candidates and the evaluator have common knowledge of the model and its parameters. The evaluator allocates grants to the applicants that generate the most favorable noisy signals. The timing is as follows:

i. Candidates observe their own type θ and decide whether to apply.
ii. The evaluator awards the available budget of grants to the applicants based on the signal realizations x.

17. As shown by Lehmann (1988), any decision maker with monotone decision preferences (Karlin and Rubin 1956) gains from a noise reduction. More generally, any decision maker with preferences in the general interval dominance ordered class introduced by Quah and Strulovici (2009) obtains a higher expected payoff state by state when σ is reduced. In addition to monotone decision problems, this preference class also encompasses single-crossing preferences (Milgrom and Shannon 1994).

18. When the noise is additive, inverting the signal distribution $y = F(\frac{x-\theta}{\sigma})$, the quantile function of the signal is $x = \theta + \sigma F^{-1}(y)$. For every percentile y, the quantile difference $[\theta + \sigma F^{-1}(y)] - [\theta + \bar{\sigma} F^{-1}(y)]$ decreases in y for $\sigma < \bar{\sigma}$. Equivalently, the quantile transform $\theta + \sigma F^{-1}(F(\frac{x-\theta}{\bar{\sigma}})) = \frac{\sigma x}{\bar{\sigma}} + (1 - \frac{\sigma}{\bar{\sigma}})\theta$ is increasing in θ for $\sigma < \bar{\sigma}$.

III.A. Fixed-Budget Equilibrium

To illustrate the logic of the model, this section considers the case with a fixed budget of grants, B. In general, equilibria have the following monotonic structure, allowing us to solve the model through a simple representation in terms of demand and supply, even though no prices are involved:

- On the supply side, the evaluator awards grants to applications with $x \geqslant \hat{x}$, because $E[\theta \mid x]$ increases in x by the MLR property.
- On the demand side, candidates with higher merit are more likely to win by MLR and thus apply for $\theta \geqslant \hat{\theta}$.

As we show, there always exists a unique fixed-budget equilibrium and this equilibrium is stable. This version of the model allows us to uncover the logic that drives the comparative statics with respect to noise: an increase in noise necessarily raises the number of applications submitted in the fixed-budget equilibrium.

1. *Application Demand: Self-Selection.* Expecting the evaluator to accept whenever the signal is above $\hat{x}$, candidates apply if their benefit from the grant times the expected probability of obtaining a grant outweighs the application cost

$$v\,[1 - F_\sigma\,(\hat{x}\,|\,\theta)] \geqslant c. \tag{4}$$

For any given acceptance standard $\hat{x}$, by the MLR property candidates apply if $\theta \geqslant \hat{\theta}$, where $\hat{\theta}$ is the marginal applicant implicitly defined by

$$1 - F_\sigma(\hat{x}\,|\,\hat{\theta}) = \frac{c}{v}, \tag{5}$$

the type whose winning probability is equal to the cost-benefit ratio.

The top panel of Figure II illustrates the signal distribution functions for the marginal type $\hat{\theta}$ and for an inframarginal type $\theta' > \hat{\theta}$. The horizontal axis corresponds to the signal realization x. Note that the distribution for θ' lies to the right of the distribution for $\hat{\theta}$, given that the MLR property implies first-order stochastic dominance. Inverting (5), the acceptance standard $\hat{x}$ that makes type $\hat{\theta}$ indifferent about applying satisfies

$$\hat{x} = F_\sigma^{-1}\left(1 - \frac{c}{v}\,|\,\hat{\theta}\right). \tag{6}$$

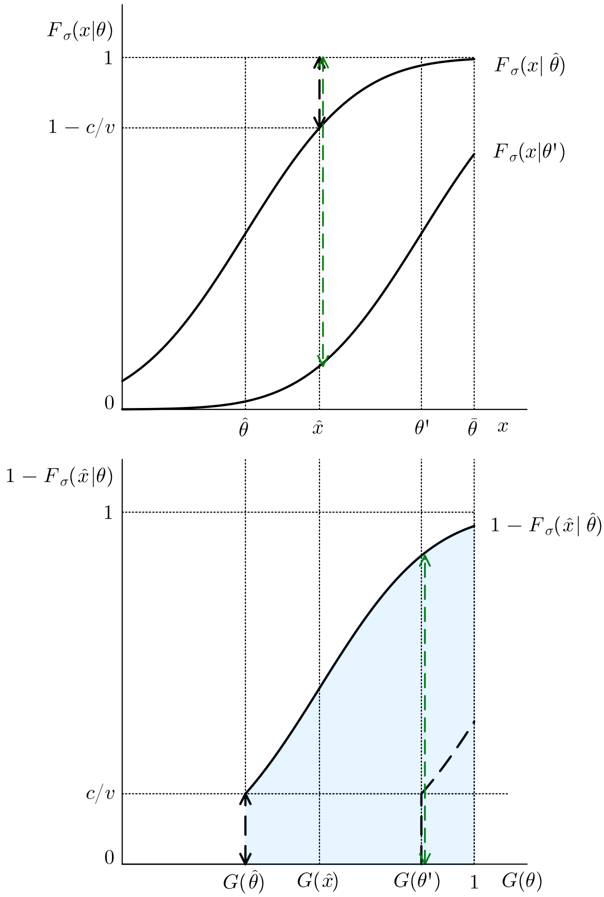

FIGURE II

Fixed-Budget Equilibrium Construction

Top panel: The black (solid) curves depict the signal distributions for $\hat{\theta}$ (curve to the left) and $\theta' > \hat{\theta}$ (curve to the right) with noise σ as a function of the signal realization x. The winning probabilities at acceptance standard $\hat{x}$ for type $\hat{\theta}$ and type θ' are marked in both panels as black (long dashed) and green (short dashed; color version available online) vertical segments with arrows. Bottom panel: The black curve is the winning probability as a function of the type percentile $G(\theta)$. According to the demand condition, the winning probability of the marginal type $\hat{\theta}$ is equal to $\frac{c}{v}$. The light blue area under the winning probability is the amount of grants awarded to all applicants.

This indifference for the marginal type pins down the demand. In the top panel, the winning probability for a given type θ can be visualized as the difference between 1 and the value of the distribution of the signal computed at $x = \hat{x}$, according to equation (5): for the marginal type $\hat{\theta}$ the winning probability is equal to $\frac{c}{v}$.

The winning probability for inframarginal type $\theta' > \hat{\theta}$ is higher than $\frac{c}{v}$, as depicted in the figure. The bottom panel of Figure II directly displays the winning probability $1 - F_\sigma(\hat{x} \mid \theta)$ as an increasing function of the agent type θ on the horizontal axis, for given acceptance standard $\hat{x}$. Thus, all inframarginal types strictly prefer to apply.

Given that the distribution function increases in the signal but decreases in the type, we can see from equation (5) that the marginal type, $\hat{\theta}(\hat{x})$, is an increasing function of the acceptance standard, $\hat{x}$. The application demand

$$a^D(\hat{x}) = 1 - G(\hat{\theta}(\hat{x})) \tag{7}$$

is then a downward-sloping function of the acceptance standard, $\hat{x}$. As the acceptance standard increases, it becomes more difficult to obtain a grant, inducing fewer candidates to apply. The marginal applicant $\hat{\theta} = G^{-1}(1-a)$ expects to obtain a grant with probability $1 - F(\hat{x} \mid G^{-1}(1-a))$. Setting the winning probability for the marginal applicant equal to the cost-benefit ratio and solving for the acceptance standard that makes the marginal applicant indifferent, we conclude:[19]

PROPOSITION 1A. (Demand) The evaluator can induce a applications by setting the acceptance standard at

$$\hat{x}^D(a) = F_\sigma^{-1}\left(1 - \frac{c}{v} \mid G^{-1}(1-a)\right). \tag{8}$$

The inverse demand is downward sloping: to induce more candidates to apply, the evaluator must reduce the acceptance standard.

2. *Grants Awarded: Evaluation.* Having derived the demand condition, the second key step of the equilibrium

19. In the special case with additive noise, $x = \theta + \sigma\varepsilon$, the marginal type is $\hat{\theta} = \hat{x} - \sigma F^{-1}(1 - \frac{c}{v})$, demand is $a^D(\hat{x}) = 1 - G(\hat{x} - \sigma F^{-1}(1 - \frac{c}{v}))$, and inverse demand is $\hat{x}^D(a) = G^{-1}(1-a) + \sigma F^{-1}(1 - \frac{c}{v})$. When information is perfect ($\sigma = 0$), the inverse demand is equal to the counterquantile function of the type distribution.

construction turns on the answer to the following question: How many grants must be awarded to induce a candidates to apply? According to the demand condition (8), by setting the acceptance standard at $\hat{x}^D(a)$, each type above the marginal, $\theta \geqslant G^{-1}(1-a)$, self-selects into applying and obtains a grant with a probability $1 - F(\hat{x}^D(a) \mid \theta)$. The grants awarded are then

$$A(a) = \int_{G^{-1}(1-a)}^{\bar{\theta}} [1 - F(\hat{x}^D(a) \mid \theta)] g(\theta)\, d\theta, \tag{9}$$

the sum of the winning probability of all applicants, weighted by their density. As applications increase, awards increase through two channels. First, the additional applicants are awarded some grants whenever they clear the acceptance standard. Second, to induce more applications, the acceptance standard $\hat{x}^D(a)$ must be reduced, thus resulting in more awards to inframarginal applicants. Overall:

PROPOSITION 1B. (Grants Awarded: Monotonicity) To induce a applicants, the evaluator must award $A(a)$ grants according to equation (9), an increasing function of a.

3. *Fixed-Budget Equilibrium.* A fixed-budget equilibrium results when the budget of grants available is equal to the budget of grants awarded, according to equation (9). As in all specifications of the model, equilibrium existence follows by the intermediate value theorem, given that the award function is continuous. An equilibrium is defined to be stable if any local perturbation leads back to the equilibrium. We have:

PROPOSITION 1C. (Fixed-Budget Equilibrium) There exists a fixed-budget equilibrium and it is unique and stable.

4. *Impact of Noise.* What is the effect of an increase in Lehmann noise to $\bar{\sigma} > \sigma$? As a first step, we show that the effect of an increase in noise on application demand $a^D(\hat{x})$, holding fixed the acceptance standard $\hat{x}$, can be positive or negative. The sign of the impact depends on whether the initial marginal type $\hat{\theta}$ benefits or is harmed by the increase in noise:

i. The top panel of Figure III illustrates an example in which the acceptance standard is above the marginal type, $\hat{x} > \hat{\theta}$. In this case, an increase in evaluation noise

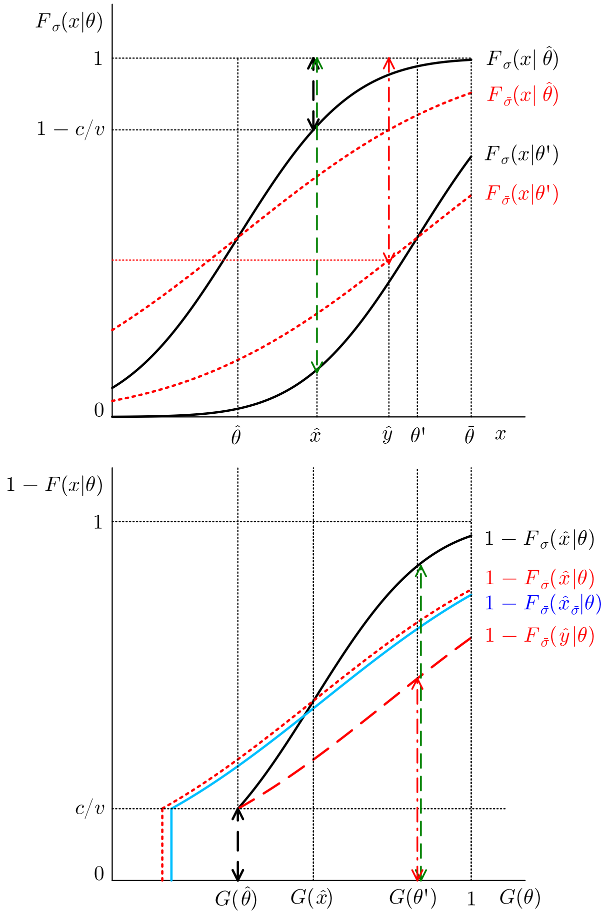

FIGURE III

Impact of Noise Increase to $\bar\sigma > \sigma$

Top panel: Impact on signal distributions. As noise increases, the signal distributions shift from the black (solid) to the red (dotted) curves. Bottom panel: Impact on winning probabilities. If the marginal type $\hat\theta$ is held constant (applications do not change), as noise increases the winning probability of inframarginal type θ' is reduced from the green (dashed) to the red (dot-dashed) segment, marked with arrows in both panels. Grants awarded under the dashed red curve are below the budget. To spend the initial budget, the marginal applicant must be reduced, as shown by the blue (in print, light-gray solid) segment, equating the area under the blue curve to the area under the black curve.

from σ to $\bar{\sigma}$ (corresponding to the dotted red distributions) benefits the initial marginal applicant $\hat{\theta}$ by raising this applicant's probability to obtain the grant, for the given acceptance standard. The new marginal applicant has a lower type, thus resulting in an increase in application demand, $a^D(\hat{x})$.

ii. If, instead, the acceptance standard were below the marginal type, $\hat{x} < \hat{\theta}$, the effect of noise would be reversed. This would be the case in the symmetric signal example in Figure III if $\frac{c}{v} > \frac{1}{2}$. In this case, noise would reduce the winning probability of the initial marginal applicant, decreasing applications for any given acceptance standard.[20]

However, once the supply side of the model is introduced, the impact of Lehmann noise on equilibrium applications becomes unambiguous. To see this, as a second step, modify the acceptance standard to restore indifference for the initial marginal type $\hat{\theta}$. To ensure that the winning probability for type $\hat{\theta}$ remains constant at the initial level, set the standard at $\hat{y}$ implicitly defined by

$$1 - F_{\bar{\sigma}}(\hat{y}|\hat{\theta}) = 1 - F_{\sigma}(\hat{x}|\hat{\theta}). \tag{10}$$

Inverting equation (10) and substituting equation (6), we obtain the explicit expression for the adjusted acceptance standard,

$$\hat{y} = F_{\bar{\sigma}}^{-1}(F_{\sigma}(\hat{x} \mid \hat{\theta}) \mid \hat{\theta}) = F_{\bar{\sigma}}^{-1}\left(1 - \frac{c}{v} \mid \hat{\theta}\right). \tag{11}$$

The top panel of Figure III illustrates the construction.

Consider now an inframarginal type $\theta' > \hat{\theta}$, who strictly prefers to apply at the initial standard $\hat{x}$. At the adjusted standard, $\hat{y}$, the winning probability for type θ' decreases provided that

$$1 - F_{\bar{\sigma}}(\hat{y} \mid \theta') < 1 - F_{\sigma}(\hat{x} \mid \theta'). \tag{12}$$

20. With additive noise, we have $\hat{x} - \hat{\theta} = \sigma F^{-1}(1 - \frac{c}{v})$, so that demand for fixed $\hat{x}$ increases or decreases in noise, $\frac{da^D}{d\sigma} \gtreqless 0$, whenever $\frac{c}{v} \lesseqgtr 1 - F(0)$. If the signal distribution is symmetric, as in the normal example used in the figure, $F(0) = \frac{1}{2}$. If $\varepsilon \geqslant 0$ we have $F(0) = 0$ so that demand always increases in noise; demand always decreases in noise if instead $\varepsilon \leqslant 0$.

We now link this condition to the Lehmann informativeness criterion. Substituting equations (11) and (6), we obtain

$$F_{\sigma}\left(F_{\sigma}^{-1}\left(1-\frac{c}{v}\,|\,\hat{\theta}\right)\,|\,\theta'\right) < F_{\bar{\sigma}}\left(F_{\bar{\sigma}}^{-1}\left(1-\frac{c}{v}\,|\,\hat{\theta}\right)\,|\,\theta'\right).$$

This condition holds under equation (3), which in turn is equivalent to signal $F_{\bar{\sigma}}$ being Lehmann-noisier than signal F_{σ} by Lemma 1 in Online Appendix A. Intuitively, an increase in noise reduces meritocracy and thus makes the winning probability for any type less responsive to merit. As seen in the bottom panel of Figure III, when the acceptance standard is adjusted to keep the initial marginal type indifferent, the winning probability of all inframarginal types is reduced, as illustrated by the shift from the black to the dashed red curve.

Weighting equation (12) by the density of the corresponding inframarginal types and summing equation (12) over all $\theta' \geqslant \hat{\theta}$, we conclude that the budget assigned will be underspent,

$$\int_{\hat{\theta}}^{\bar{\theta}} [1 - F_{\bar{\sigma}}(\hat{y}\,|\,\theta)]\, g(\theta)\, d\theta < \int_{\hat{\theta}}^{\bar{\theta}} [1 - F_{\sigma}(\hat{x}\,|\,\theta)]\, g(\theta)\, d\theta,$$

whenever we retain indifference by the initial marginal type.

PROPOSITION 1D. (Impact of Noise on Award Function) As noise σ in the evaluator's signal increases, fewer grants are awarded for any given level of applications.

Given a fixed budget B, at the higher noise $\bar{\sigma}$ more applications from agents with types below the initial $\hat{\theta}$ must be encouraged in the new equilibrium by lowering the acceptance standard below $\hat{y}$. Thus, we obtain our keystone comparative statics:

PROPOSITION 1E. (Impact of Noise on Fixed-Budget Equilibrium Applications) As noise σ in the evaluator's signal increases, fixed-budget equilibrium applications increase.

The remainder of the article shows that this comparative statics result holds more generally—and is actually strengthened—when the budget allocated to a field increases with applications. Before proceeding, we step back and prod the robustness of this result to the simplifying assumption that candidates have perfect information about their merit.

5. *Noisy Self-Selection.* Does evaluation noise increase applications also when candidates have a noisy, rather than perfect,

signal t about their type θ? For the case with noisy bilateral information, we can leverage the quantile function approach to easily extend the result once we restrict it to parametric signals, for which we can prove the Lehmann (1988) property. For example, suppose types are normally distributed, $\theta \sim N(0, 1)$, and the evaluator, as well as the candidates, observe normal and conditionally independent signals, $x \,|\, \theta \sim N(\theta, \sigma^2)$ and $t \,|\, \theta \sim N(\theta, \tau^2)$, respectively. To decide whether to apply, candidates must now forecast what their type is likely to be. Upon observing signal t, a candidate's updated belief about the type is $\theta|t \sim N(\frac{1}{1+\tau^2}t, \frac{\tau^2}{1+\tau^2})$. Candidates with higher signals are more likely to believe their type is high. Knowing that the evaluator observes a noisy signal, $x \,|\, \theta \sim N(\theta, \sigma^2)$, the candidate's belief about the signal the evaluator observes is $x|t \sim N(\frac{1}{1+\tau^2}t, \sigma^2 + \frac{\tau^2}{1+\tau^2})$, so that an increase in σ reduces Lehmann (1988) information. Thus, exploiting the general argument presented above, an increase in the evaluator noise σ makes the winning probability less responsive to the candidate's signal about the type. Applications increase for any given awards budget, as in the baseline model. In addition, we can establish that an increase in candidate noise τ also reduces Lehmann's (1988) information and thus reduces applications.

III.B. Partial Equilibrium with Subproportional Budget Allocation

This section considers a single-field model where the budget of grants $B(a)$ depends on applications, where a is the fraction of applicants within the unit size population of candidates in the field. We restrict attention to the budget rules that are (weakly) increasing and subproportional

$$\frac{\partial}{\partial a}\frac{B(a)}{a} \leqslant 0, \tag{13}$$

that is, the grant budget per application (weakly) decreases in applications.[21] Graphically, the segment that connects any point (a, $B(a)$) in the graph to the origin (0,0) lies entirely below the graph itself. Equivalently, the rays of the function become less steep as a increases so that none of the area below the graph of the function is hidden from an observer at the origin by the graph itself. Intuitively, inequality (13) relaxes concavity by requiring the

21. This is the opposite of a differentiable version of star-shaped, as defined by Marshall and Olkin (2007, 690–691).

function's average, rather than the derivative, to decrease. The subproportional budget case encompasses a fixed budget, $B(a) = B$, as well as the case with constant payline, $B(a) = pa$, where the fraction of grants is proportional to applications. This formulation allows us to deal with a partial equilibrium version of the full model where the payline $p(a)$ decreases in applications in a field, holding fixed the number of applications in all other fields.

1. *Shape of the Award Function.* The characterization of the equilibrium in terms of uniqueness, stability, and comparative statics hinges on the shape of the award function (9), which gives the grant awards necessary to induce a candidates to apply. Notice that the award function in the example depicted in the bottom panel of Figure II is superproportional

$$\frac{\partial}{\partial a}\frac{A\,(a)}{a} \geqslant 0,$$

as illustrated in the top panel of Figure IV. This is the opposite of condition (13) that we imposed for the budget. From now on, we restrict attention to signals with additive noise. In this case, the superproportionality of the award function hinges on the monotonicity of the hazard rate of the type distribution:

PROPOSITION 2A. (Grant Awarded: Shape) The award function $A(a)$ is superproportional, linear, or subproportional if the type distribution G has respectively an increasing, constant, or decreasing hazard rate.

To understand the logic behind this central result, rewrite the integrand in equation (9) as a function of the type percentile $t = G(\theta)$

$$A(a) = \int_{1-a}^{1} [1 - F(\hat{x}^D(a) \,|\, G^{-1}(t))]dt.$$

Thanks to this change of variable, the budget necessary to induce a applications can be visualized as the area below the winning probability curve, as represented in the bottom panel of Figure II. Equivalently, we can express this area as a rectangle with a base that spans the integration segment (of length a) and height equal to the average winning probability under a applications, $\frac{A(a)}{a}$. The average winning probability when a fraction a of the population applies is precisely the average of the budget necessary to induce a applications. Graphically, the average winning probability is the

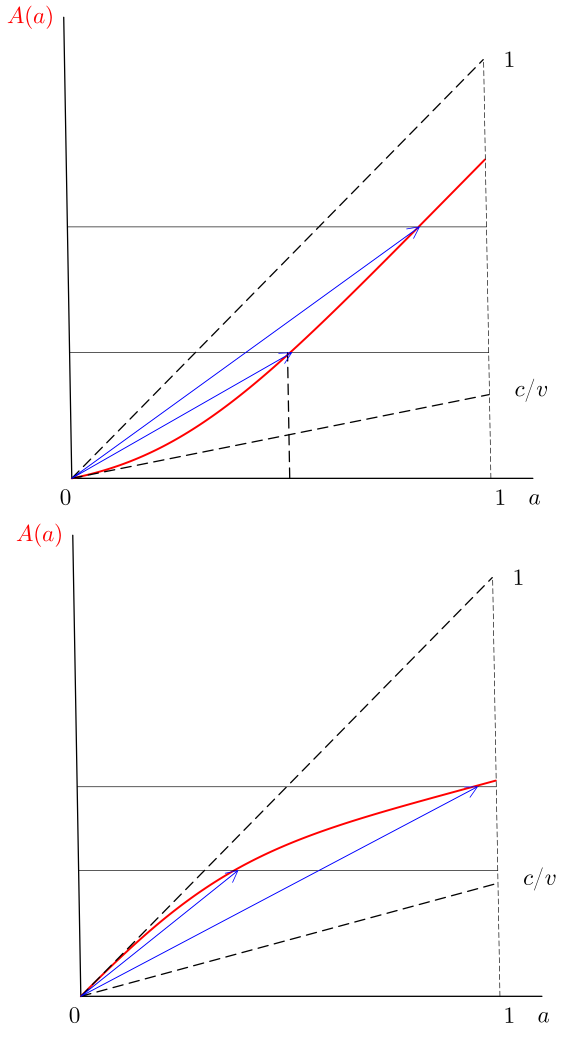

FIGURE IV

Award Function

Top panel: Superproportional award function, with increasing rays. Bottom panel: Subproportional award function, with decreasing rays.

slope of the segment connecting $(a, A(a))$ to the origin in the two panels of Figure IV. According to the claim, if applications increase, the average winning probability increases if and only if the hazard rate of the type distribution G is increasing.

As applications increase, the acceptance standard must be reduced to attract more applications. Thus, the winning probability of all inframarginal applicants must increase.[22] The effect of the increase in applications on the average winning probability among all applicants, however, hinges on the relative increase in winning probabilities for applicants at different percentiles of the type distribution.

The award function is superproportional when stronger applicants, with types at higher percentiles, absorb a larger fraction of grants—that is, when the increase in the rent that inframarginal types obtain as they win with higher probability increases in the type percentile. Intuitively, if the type distribution has an increasing hazard rate (or decreasing hazard rate), the distance between the average type of the inframarginal applicants and the type of the marginal applicant, $E(\theta|\theta > \hat{\theta}) - \hat{\theta}$, decreases (or increases) in the type of the marginal applicant

$$\frac{\partial}{\partial \hat{\theta}}[E(\theta|\theta > \hat{\theta}) - \hat{\theta}] < 0 \text{ (or } > 0),$$

see Bagnoli and Bergstrom (2005), Theorem 6.[23] Under an increasing hazard rate, inframarginal applicants on average become stronger relative to the marginal applicant, when the marginal applicant is reduced as a result of an increase in applications. When the signal is additive, an increase in the difference $E(\theta|\theta > \hat{\theta}) - \hat{\theta}$ translates into a higher gap between the average winning probability of the inframarginal applicants and the winning probability of the marginal applicant. Given that along the demand curve the winning probability of the marginal applicant

22. In the bottom panel of Figure II, for example, if the acceptance standard were lowered so as to reduce the marginal type from θ' to $\hat{\theta} < \theta'$, the winning probability for the inframarginal types would increase from the dashed curve to the continuous curve.

23. Increasing (or decreasing) hazard rate of the type distribution is equivalent to log-concavity (or log-convexity) of the countercumulative distribution $1 - G(\theta)$, which implies log-concavity (or log-convexity) of the right-hand integral $H(\theta) = \int_{\theta}^{\bar{\theta}} [1 - G(\tilde{\theta})]d\tilde{\theta}$, which in turn is equivalent to the fact that the residual expectation $E[\theta - \hat{\theta}|\theta \geqslant \hat{\theta}]$ is decreasing (or increasing) in $\hat{\theta}$.

is fixed at $\frac{c}{v}$ by construction, in equilibrium the average winning probability of the inframarginal applicants $\frac{A(a)}{a}$ increases in a.

The opposite conclusion holds if the type distribution has a decreasing hazard rate. Reversing the logic, the budget of awards needed to incentivize additional applicants increases less than proportionally with the number of applications. Under a decreasing hazard rate, stronger applicants absorb relatively fewer incremental grants than weaker applicants. As the winning probability rises, the proportion of grants awarded to applicants with higher types decreases. Given that relatively fewer applicants win with higher probability, the average winning probability is reduced as applications increase.

In the borderline case with constant hazard rate, the type distribution is exponential. As applications increase, the distance between the type of the marginal applicant and the average type of all inframarginal applicants remains constant. Suppose the application level is a, resulting in an average winning probability $\frac{A(a)}{a}$. To increase applications to $a' > a$, the acceptance standard must be reduced—as a result, the winning probability of all inframarginal applicants goes up. As applications increase, the number of applicants who expect to win a grant with a probability above any given level rises, but with a constant hazard rate it remains constant as a fraction of the number of applications. As a consequence, the average winning probability $\frac{A(a)}{a} = \frac{A(a')}{a'}$ remains constant.[24]

2. *Partial Equilibrium.* As illustrated in the top panels of Figures IV and V, with an increasing hazard rate the award function is superproportional and crosses once and from below the subproportional budget function for an interior $a \in (0, 1)$, provided that $B'(0) > A'(0)$ and $B(1) < A(1)$.[25] This equilibrium is stable, given that a small increase (or decrease) in a above (or

24. The precise condition derived in Proposition 2 hinges on the assumption of additive noise. Similar results can be derived for more general signal distributions. For example, if the signal follows a Kumaraswamy distribution, $F(x\,|\,\theta) = 1 - (1 - x^{\theta})^{b}$ with parameter $b > 0$ (resulting in less precise evaluation for higher types), it can be shown that the award function is superproportional, linear, or subproportional whenever the elasticity of the type distribution is increasing, constant, or decreasing. When the type distribution G has positive support, the borderline case is given by the Pareto distribution.

25. These two conditions are rather natural. If the hazard rate of the type distribution is unbounded, we have $A'(0) = \frac{c}{v}$, so that $B'(0) > A'(0)$ avoids the

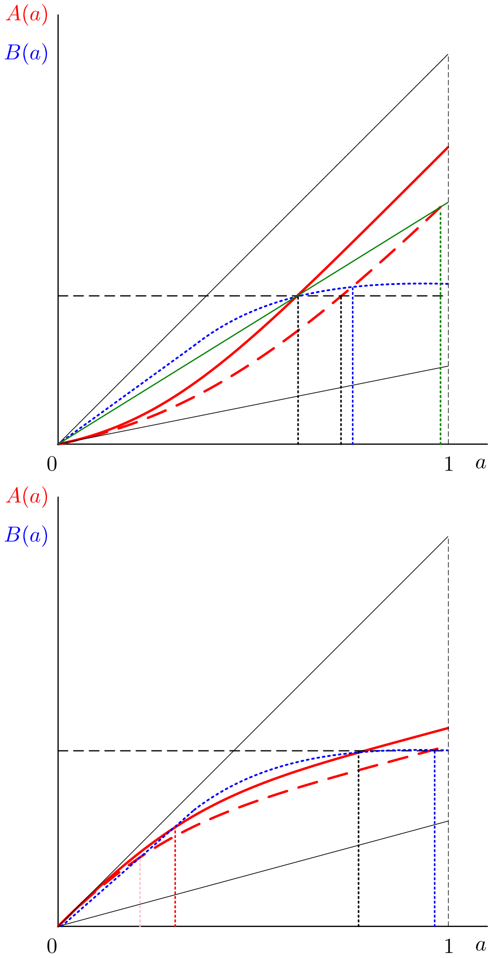

FIGURE V

Comparative Statics

Top panel: A superproportional award function (red heavy solid) crosses a subproportional budget (blue dotted) only once from below, resulting in a unique stable equilibrium. As noise increases, the award function shifts to the right to the red dashed curve, resulting in a larger increase in applications than under a fixed budget (dashed black horizontal segment), but smaller than under constant payline (straight thin green segment). Bottom panel: Multiple equilibria are possible when the award function is subproportional. Here there is an unraveling stable equilibrium, an unstable equilibrium with intermediate applications, and a stable equilibrium with high applications. As noise increases, applications increase in the interior stable equilibrium, but decrease in the unstable equilibrium.

below) the equilibrium level results in an increase (or decrease) in grants awarded above (or below) the budget, thus inducing an adjustment back to the equilibrium:[26]

PROPOSITION 2B. (Partial Equilibrium) If the type distribution has an increasing hazard rate and the budget rule is subproportional, there is a unique partial equilibrium and this equilibrium is stable.

The bottom panels of Figures IV and V illustrate subproportional award functions resulting when the type distribution has a decreasing hazard rate. To understand the role of the hazard rate condition on the equilibrium, consider the special case with a proportional budget, $B(a) = pa$, where $p > 0$ represents the grants available per application. In this case with constant payline, when the type distribution has a decreasing hazard rate, the unique stable equilibrium is always at the corner. If $A'(0) > p$, unraveling $a = 0$ results in the unique stable equilibrium; if instead $A'(0) < p$, all agents apply $a = 1$ in the unique stable equilibrium.

3. *Stimulus Bill.* Proposition 2B has an immediate implication for the effect of an (anticipated) increase in the budget on the success rate—also known as payline—the widely reported fraction of successful applications:

PROPOSITION 2C. (Impact of Budget on Success Rate) If the type distribution has an increasing, constant, or decreasing hazard rate, the equilibrium success rate respectively increases, is constant, or decreases in the budget.

This prediction can be confronted with the outcome of the increase in the NIH budget after the 2009 stimulus bill. As part of the stimulus bill introduced by the U.S. Congress in 2009 in the aftermath of the great financial crisis, the American Recovery and Reinvestment Act (ARRA) allocated an additional \$8.97 billion to extramural research grants at the NIH in two parts:

trivial case in which the budget is so scarce that nobody applies. Condition $B(1) < A(1)$ imposes that the budget is too scarce to accommodate all applications.

26. When $B'(0) < A'(0)$, there is a stable corner equilibrium with unraveling $a = 0$, in which nobody applies. When $B(1) > A(1)$, there is a stable corner equilibrium $a = 1$ in which all agents apply. In all cases, the equilibrium is unique and stable.

- Part of the funds (19.3%) of the total ARRA budget appropriated to the NIH was allocated to “not ARRA solicited” applications that had been previously submitted and reviewed in recent evaluation cycles but were marginally rejected. Park, Lee, and Kim (2015) empirically document that “not ARRA solicited” applications resulted in fewer high-impact articles than did regular projects.
- The remainder of the funding bonanza was set aside to increase the budget for “ARRA solicited” grant competitions. In this case, potential applicants were informed of the larger budget. A second fact, documented by Stephan (2012, 145), is that such applications increased so much that the success rate decreased.

These two facts can be rationalized in our model. First, the budget allocated to “not ARRA solicited” applications corresponds to an unanticipated increase in the budget. Compared to the prepolicy equilibrium, the policy change shifts the supply to the right. Holding fixed the number of applications at the equilibrium level a^B for the initial budget B, the model predicts that the applications funded as a result of the increase in the budget to $B' > B$ are of lower quality, by the MLRP of the signal (2).

Second, consider the impact of an anticipated increase in the budget on the success rate. According to Proposition 2D, applications must increase more than proportionally with the budget for the success rate to decrease as the budget increases—and this occurs in equilibrium if and only if the type distribution has a decreasing hazard rate. This is exactly what happened as a result of the “ARRA solicited” part of budget increase in 2009. This observation is consistent with a type distribution with decreasing hazard rate at the top, as is natural to expect for the talent of scientists and artists; see Seglen (1992) and Caves (2000).[27]

27. Distributions with a decreasing hazard rate are more right-skewed than the exponential distribution. They can be obtained by stretching an exponential distribution toward the top tail through a convex transformation. A distribution has a decreasing hazard rate whenever it is larger than the exponential distribution in van Zwet’s (1964) convex transform order. Given two distributions G and H, van Zwet (1964) defines G to be smaller than H in the convex transform order, denoted $G \prec_c H$, whenever $H^{-1}(G(\cdot))$ is convex. As shown by van Zwet (1964), a distribution G with an increasing (or decreasing) hazard rate can be obtained through an increasing and concave (or convex) transformation $G^{-1}(G_{\text{Exp}}(\cdot))$ of a random variable with an exponential distribution. To gain intuition, visualize the random variable G^{-1} on the vertical axis as an increasing transformation of an

4. *The Impact of Noise on Applications.* We return to our headline comparative statics with respect to evaluation noise σ. Recall from Proposition 1D that an increase in noise σ shifts down the award function. Given that the budget function is increasing and subproportional, if the award function is superproportional, as in the top panel of Figure V, we conclude that applications in the unique and stable equilibrium increase in σ more than under a constant budget.[28] More generally, the following headline comparative statics result holds for all stable equilibria:

PROPOSITION 2D. (Partial Equilibrium) As noise σ in the evaluator's signal increases, the application level a increases in any stable partial equilibrium.

5. *Grading on a Curve and the Paradox of Relative Evaluation.* Turning to an extreme version of this comparative statics result, consider the outcome resulting when the evaluation is based on a perfect signal without noise, $\sigma = 0$. With fixed budget B, the most efficient allocation results: the best a agents apply and obtain grants with probability 1.

What if, instead, grants can only be awarded to a fraction $p < 1$ of applicants? This case corresponds to grading on a curve with binary classification, with awards for the best $p < 1$ participants. Equivalently, suppose that the budget is proportional to applications, $B = pa$, with a constant payline, corresponding to a small panel under the NIH payline system or under the ERC's PA rule. Given any acceptance standard x, with perfect information, all applicants with $\theta \geqslant x$ anticipate that they will succeed and thus apply to secure $v > c$. However, only a fraction $p < 1$ of these applicants can succeed. Thus, if $a > 0$, a fraction $1 - p$ of applicants cannot succeed. But applicants with types below the $1 - p$ quantile of the conditional type distribution among applicants, having perfect information and thus anticipating that they will not succeed, strictly prefer not to apply and save the application cost. The process continues until we obtain that the constant payline equilibrium for $p < 1$ with perfect information ($\sigma = 0$) always

exponential random variable G_{Exp}^{-1} on the horizontal axis through a Q–Q plot. Concavity (or convexity) of $G^{-1}(G_{\text{Exp}}(\cdot))$ contracts (or stretches) the top tail and makes it thinner (or thicker) than the top tail of an exponential.

28. The comparative statics result holds strictly for interior equilibria, but weakly for corner equilibria.

unravels: the unique (stable) equilibrium features no participation, $a = 0$.[29]

This unraveling logic highlights how grading on the curve, if perfect, destroys participation incentives. More generally, this logic immediately extends to any budget rule with the property that $\frac{B(a)}{a} < 1$ for any a:

PROPOSITION 2E. (Unraveling) If the evaluation is based on perfect information, $\sigma = 0$, and $B(a) < a$, in the unique partial equilibrium no candidate applies.

The result follows immediately from the observation that the award function without noise is $A(a) = a$. Unraveling starts at the bottom of the distribution, where applicants pull out, anticipating that they stand no chance of winning, thus reducing the budget available for the top applicants. In the end, symmetric information destroys incentives for costly participation in this nonmarket environment, turning on its head Akerlof's (1970) classic insight that asymmetric information reduces incentives for market participation.

6. *Unraveling with Noisy Information.* It is worth stressing that although perfect information is sufficient, it is not at all necessary for unraveling. Well beyond the case with perfect information, unraveling results provided that $B(0) = 0$ and $B'(0) < A'(0)$ where $A'(0) := \lim_{a\to 0^+} A'(a)$. For example, we verified that with uniform additive noise, $F(\varepsilon) = \frac{1}{2} + \varepsilon$, whenever $B(0) = 0$ with $B'(0) < 1$ (and thus for any constant payline $p < 1$):[30]

- For type distributions with a bounded hazard rate, $\lim_{\theta\to\infty} \frac{g(\theta)}{1-G(\theta)} < \infty$, such as logistic $G(\theta) = (1 + \exp(-\frac{\theta-\mu}{s}))^{-1}$ there is a stable equilibrium with unraveling, $a = 0$, not only in the absence of noise but also when noise is limited, $\sigma \leqslant \tilde{\sigma}$ with $\tilde{\sigma} > 0$.[31]

29. Or, equivalently, only the highest type $\overline{\theta}$ (measure-zero) applies and is awarded a fraction p of the grant.

30. The derivative of the award function is then $A'(a) = \frac{c}{v} + \frac{G(G^{-1}(1-a)+(1-\frac{c}{v})\sigma)-(1-a)}{\sigma g(G^{-1}(1-a))}$.

31. With logistic types the award function is $A(a) = a\frac{c}{v} - \frac{s}{\sigma}\ln(1 - a + a\exp(-(1-\frac{c}{v})\frac{\sigma}{s}))$. From $A'(0) = \frac{c}{v} + \frac{s}{\sigma}[1 - \exp(-(1-\frac{c}{v})\frac{\sigma}{s})]$, we have $\frac{\partial A'(0)}{\partial\sigma} < 0$, $\lim_{\sigma\to 0} A'(0) = 1$ and $\lim_{\sigma\to\infty} A'(0) = \frac{c}{v}$, so that $\tilde{\sigma}$ uniquely solves $\frac{c}{v} + \frac{s}{\tilde{\sigma}}[1 - \exp(-(1-\frac{c}{v})\frac{\tilde{\sigma}}{s})] = B'(0)$. E.g., for $\frac{c}{v} = 0.2$ and $p = 0.3$, we have $\tilde{\sigma} = 10s$.

- For type distributions with a vanishing hazard rate, $\lim_{\theta\to\infty} \frac{g(\theta)}{1-G(\theta)} = 0$, such as Pareto $G(\theta) = 1 - \frac{1}{\theta}$ with support $[1, \infty)$ there is always an equilibrium with unraveling, for any level of noise.[32]

7. *Equilibrium Multiplicity.* When the award function is not strictly superproportional (at least on a subinterval) multiple equilibria possibly arise. The borderline case with equilibrium multiplicity features exponentially distributed types, $G(\theta) = 1 - \exp(-\lambda\theta)$, with constant hazard rate λ. The award function is then linear in a, with slope increasing in $\frac{c}{v}$ and decreasing in λ and in σ.[33] For any given payline $p > \frac{c}{v}$, there is a threshold level of noise $\tilde{\sigma}$ at which there is a continuum of equilibria for any $a \in [0, 1]$, for $\sigma < \tilde{\sigma}$ there is a unique equilibrium with unraveling, and for $\sigma > \tilde{\sigma}$ there is a unique equilibrium with $a = 1$.[34]

When the type distribution has a strictly decreasing hazard rate, the resulting subproportional award function can cross multiple times with a subproportional budget, as illustrated in the bottom panel of Figure V. The headline comparative statics from Proposition 2B applies for all stable equilibria, also when they are multiple. Given that the award and budget functions are both continuous, equilibria alternate in terms of stability. By Samuelson's (1947) correspondence principle, the sign of the comparative statics is reversed for unstable equilibria.[35]

IV. Grantmaking across Fields

We turn to the problem of grant allocation across fields $i = 1, ..., N$, each populated by a continuum of candidates

32. With Pareto types the award function is equal to $A(a) = a\frac{c}{v} + \frac{1}{\sigma}\ln(1 + (1 - \frac{c}{v})\sigma a)$, with $A'(0) = 1$ for all $\sigma > 0$. For any level of noise σ, unraveling, $a = 0$, is a stable equilibrium.

33. For example, if the signal is uniform, $F(\varepsilon) = \frac{1}{2} + \varepsilon$, the award function is $A(a) = a[\frac{c}{v} + \frac{1-\exp(-\sigma\lambda(1-\frac{c}{v}))}{\sigma\lambda}]$. If the signal is exponential, $F(\varepsilon) = 1 - \exp(-\varepsilon)$, the award function is $A(a) = \frac{a}{1-\lambda\sigma}[(\frac{c}{v})^{\lambda\sigma} - \frac{\lambda\sigma c}{v}]$.

34. For $\gamma = 0.2$ and $p = 0.3$, with a uniform signal we have $\tilde{\sigma} = 10\lambda$.

35. For an analytical example, with Pareto-distributed types as in the last bullet point, with constant payline $p < 1$, in addition to the unraveling equilibrium, for $\sigma > \hat{\sigma}$ where $\hat{\sigma} < \infty$ is the unique solution of $\frac{1}{\sigma}\ln(1 + (1 - \frac{c}{v})\sigma) = p - \frac{c}{v}$, there is also a second stable equilibrium in which everyone applies, $a = 1$, as well as an intermediate unstable equilibrium with $a \in (0, 1)$. For example, for $\frac{c}{v} = 0.2$ and $p = 0.3$, we have $\hat{\sigma} = 33.15$.

representing the pool of potential applicants. Field i is characterized by specific parameters, such as type distribution G_i, signal noise distribution F_i, noise dispersion σ_i, application cost c_i, and private benefit v_i from obtaining a grant.[36] As in the baseline model, candidates are atomistic and thus do not consider their application decision's effect on the acceptance standard. In each field, the evaluator (think of the review panel) allocates to field i a budget $B_i(a_1, ..., a_N)$ dependent on the applications submitted in all fields. In each field, grants are awarded to applications with the highest expected merit in the field.

Online Appendix B characterizes the equilibria in the model with multiple fields, with particular attention to budget rules that satisfy a multidimensional generalization of subproportionality (13), condition SPA. As shown in Proposition 4, quasi-proportional budget allocation rules

$$\text{(QPA)} \qquad B_i = \frac{a_i^{\varrho_i}}{\sum_{j=1}^{N} a_j^{\varrho_j}} B,$$

with proportionality coefficients $\varrho_i \in [0, 1]$ satisfy SPA. This class encompasses the PA rule used by the ERC, NIH, and Canadian research funding organizations (for $\varrho_i = 1$ for all i) and the fixed budget rule adopted by the NSF as well as by the U.K. and Australian agencies (for $\varrho_i = 0$ for all i), but more generally allows for field-specific budget responsiveness ϱ_i.

If we combine subproportionality with an increasing hazard rate, we obtain a unique stable equilibrium that preserves the comparative statics we derived for the partial equilibrium—applications increase in own noise—and reverses it for other fields—applications decrease in noise in other fields:

36. The model can be easily extended to allow for fields to have different sizes, n_i, and for the individual budget, q_i, that each applicant can request to vary across fields so that if fraction a_i of candidates apply in the field i the total funds requested in the field are $n_i q_i a_i$. In practice, grant calls typically set upper bounds to the size of the award applicants can ask, sometimes depending on the applicant's career stage. The ERC sets the maximum allowed awards at the same level for all fields. Given that almost all applicants request (and successful applicants are awarded) approximately the maximum allowed, we do not model the individual choice of the amount by the applicant. In the more general case where grant applicants request awards of different sizes, panel i selects the projects with the highest score to distribute the fraction $100 \times p$ of the total funds applied for in field i.

PROPOSITION 3A (Unique General Equilibrium). If the type distributions in every field have an increasing hazard rate (IHR) and the budget rule is subproportional, SPA, the general equilibrium (i) is unique, (ii) stable, and (iii) satisfies the comparative statics that an increase in noise in a field i increases applications in that field

$$\frac{da_i^E}{d\sigma_i} \geqslant 0, \tag{14}$$

and decreases applications in any other field j

$$\frac{da_j^E}{d\sigma_i} \leqslant 0. \tag{15}$$

To understand this result, note that by Proposition 1E, for any given budget size, noisier fields tend to attract more applications. As the budget in a field increases in applications, the increased number of applications results in an increase in the budget, which in turn induces a further increase in applications. If applications increase less than proportionally with the budget, as is the case when the type distribution has an increasing hazard rate, and the budget is subproportional in applications, the process converges to a unique interior equilibrium that features more applications in the noisier field and fewer applications in the other fields.

IV.A. *Multiple Equilibria*

For type distributions without an increasing hazard rate, multiple equilibria become possible. In the borderline case with exponentially distributed types in all fields, an extreme version of the paradox of relative evaluation arises: under PA of a budget $B < 1$, the entire budget is allocated to the field with the noisiest evaluation and all other fields unravel, even if their noise is infinitesimally lower. More generally, Online Appendix B shows that the comparative statics for all stable equilibria remain well behaved for symmetric budget rules such as PA:

PROPOSITION 3B (Multiple General Equilibria). For general type distributions, in any stable general equilibrium under PA applications in a field increase in the noise of that field, inequality (14).

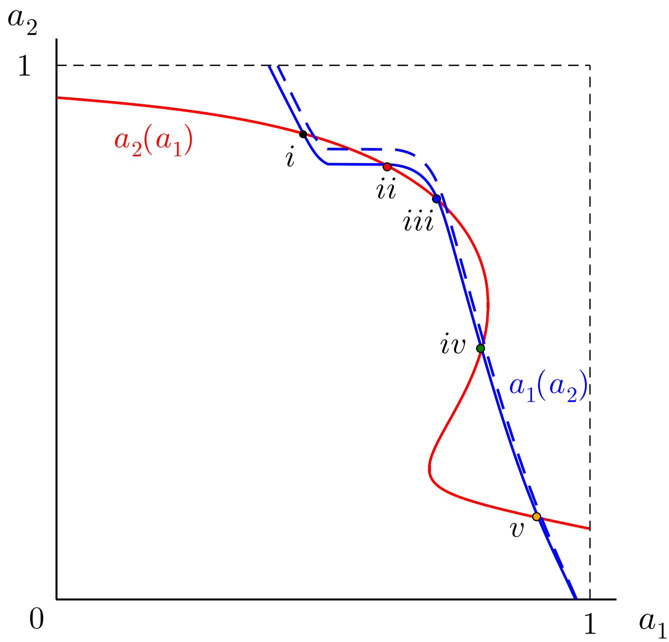

FIGURE VI

Construction of General Equilibria

The general equilibria are at the crossing of the partial equilibrium correspondences for two fields: $a_j(a_i)$ represents the set of partial equilibrium applications in field j as a function of the application level in field i. An increase in σ_1 shifts $a_1(a_2)$ to the right to the dashed blue curve. All stable equilibria, here (i), (iii), and (v) satisfy inequalities (14) and (15).

Figure VI illustrates the construction of the general equilibrium and the logic of the comparative statics result with $N = 2$ fields in an example featuring multiple equilibria. In each field, types follow a mixture of two normal distributions with a nonmonotonic hazard rate (increasing for low types, decreasing for intermediate types, and increasing again for high types). To understand the shape of field 2's partial equilibrium correspondence (the red curve in Figure VI) $a_2(a_1)$, given applications a_1 in field 1, note that the level of applications in field 1 affects the equilibrium in field 2 only through the budget function $B_2(a_1, a_2)$, which decreases in a_1. As can be seen from Figure V, bottom panel, the budget reduction created by the increase in a_1 results in a decrease in applications at any stable partial equilibrium—and in an increase in applications at any unstable partial equilibrium. The two decreasing portions of the partial equilibrium correspondence $a_2(a_1)$ in Figure VI depict stable partial equilibria, while the increasing portion depicts an unstable interior partial equilibrium. A similar construction applies to field 1's partial equilibrium correspondence $a_1(a_2)$ (the blue curve in Figure VI).

The points of intersection of the partial equilibrium correspondences $a_2(a_1)$ and $a_1(a_2)$ are general equilibria. In this example, there are five general interior equilibria, marked by colored dots in the figure. A general equilibrium is stable when $a_2(a_1)$ crosses $a_1(a_2)$ from below at points at which both $a_2(a_1)$ and $a_1(a_2)$ are downward-sloping. Here, general equilibria (i), (iii), and (v) are stable, whereas (ii) is general-equilibrium unstable ($a_2(a_1)$ crosses $a_1(a_2)$ from above) and (iv) is partial equilibrium unstable ($a_2(a_1)$ is upward-sloping at the crossing). As a result of an increase in noise in field 1, by Proposition 2D, $a_1(a_2)$ shifts to the right (dashed blue curve). All stable equilibria satisfy both own and cross-comparative statics, inequalities (14) and (15), given that at a stable equilibrium $a_2(a_1)$ slopes down and is flatter than the downward-sloping $a_1(a_2)$.

V. Empirical Validation: The 2014 ERC Funding Reform

This section exploits the natural experiment of the 2014 reform of the ERC funding rules to test and quantify our theory's central prediction about the impact of noise on applications and budget shares. Figure VII explains the different steps of the analysis. We first need to compute the evaluation noise in each ERC panel (step 1) to quantify the effect of evaluation noise on the ERC grant applications (step 2). We detail these steps after explaining the institutional background regarding the ERC funding.

V.A. *Institutional Background: ERC Funding*

The ERC funding scheme was set up in 2007 by the European Union and has funded over 10,000 researchers across all research fields with a budget of about €1.7 billion per year.[37] Before 2014, the ERC used to set the budget for each of three disciplinary domains from the top down, at about 39% for LS, 44% for PE, and 17% for SH. In each domain, the budget was allocated to panels

37. The ERC has annual calls for three separate levels of seniority: starting, consolidator, and advanced grants.

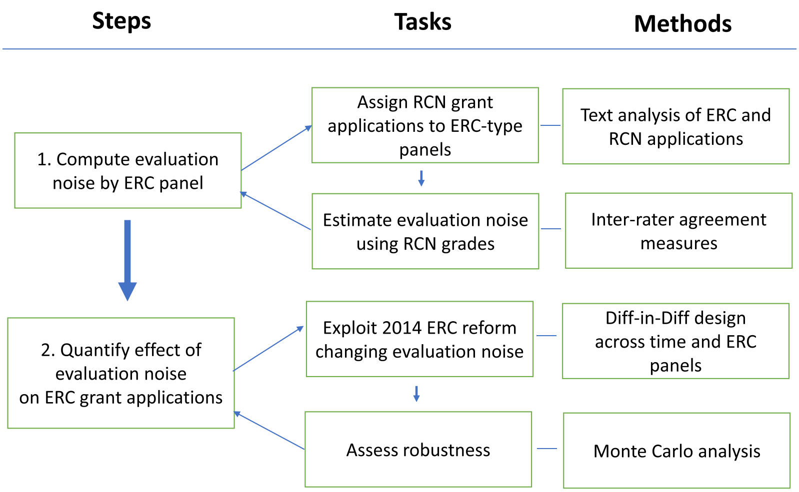

FIGURE VII

Steps in the Empirical Analysis

from the bottom up, in proportion to the budgetary demand by proposals submitted to the panels in the same domain according to PA. From 2015 the ERC started allocating funds proportionally across all panels, making each panel's budget dependent on the applications to panels belonging to the other two domains as well.[38] As shown in Figure I, the relative budgets of the three domains were stable until the reform but started to diverge from 2015, with a sharp decline in the budget share devoted to LS and an increase for SH.

The reform also had consequences within domains, as shown in Figure VIII. Within SH, panels such as SH3 (environmental studies, geography, and demography), SH5 (cultural studies), and SH6 (history) saw an increase in their share of the budget. In contrast, the relative budget of SH1 (economics, finance, and management) remained relatively constant. In LS, several basic

38. Before the reform, domain budgets $\bar{B}_{\mathrm{PE}}, \bar{B}_{\mathrm{LS}}, \bar{B}_{\mathrm{SH}}$ were fixed and then allocated to each panel i proportionally in each domain, resulting in $B_i = \frac{a_i \bar{B}_{d(i)}}{\sum_{j \in G_{d(i)}} a_j}$, where $d(i) \in \{\mathrm{PE, LS, SH}\}$ is the domain of panel i and $G_{d(i)}$ is the set of panels in the same domain as panel i. After the reform, the budget allocation to panels became proportional within and across domains, $B_i = \frac{a_i B}{\sum_{j \in G_{All}} a_j}$, where G_{All} contains all the panels in *PE*, *LS*, and *SH*.

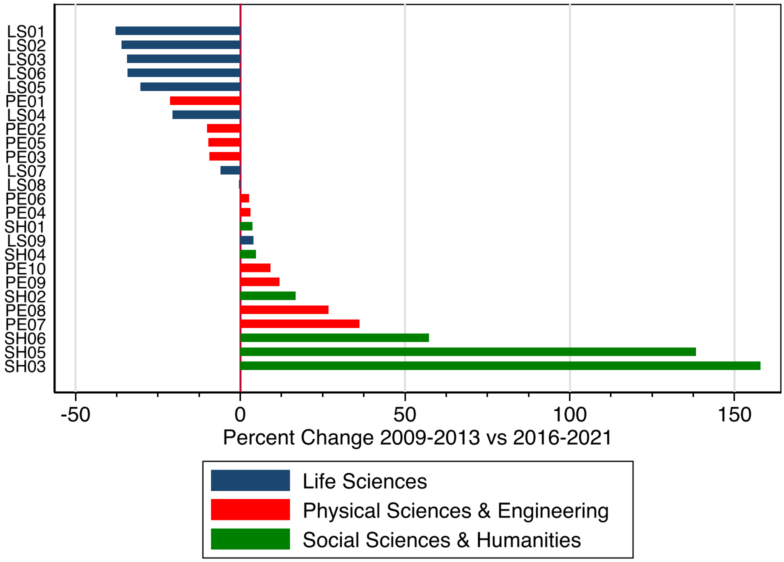

FIGURE VIII

Budget Shares in ERC Funding by Disciplinary Domain

This figure shows the relative change in funding for each ERC panel for 2009–2013 compared to 2016–2021, leaving out the years around the 2014 reform. LS: Life Sciences, PE: Physical Sciences and Engineering, SH: Social Sciences and Humanities. LS01 covers molecular biology, biochemistry, structural biology, and molecular biophysics. SH03 covers demography, sociology, anthropology, education, and communication. *Source:* ERC data.

research panels ranging from LS1 (molecular biology) to LS5 (neurosciences) saw a sustained decline, whereas the budget share of more applied panels like LS9 (nonmedical biotechnology) increased by little. In the empirical analysis, we argue that these changes can be attributed to how the budget reform leverages the effect of evaluation noise on application incentives across diverse panels.

V.B. Econometric Specification

Proposition 3 relates the number of applications a_{ist} to evaluation noise, σ_i, where the indices i, s, and t represent the panel, the seniority of the grant call, and the year, respectively. We exploit the ERC reform to inform the theory (step 2 in Figure VII) by associating to each panel a panel group G_{it} within which budget allocations are made in that year. We define N_{it} as the number of panels in group G_{it}. Group membership is changing over time due to the reform of the ERC funding. Before the reform, panels belonging to the same domain were competing for funds only with other panels belonging to the same domain; thus, panels are assigned to three groups depending on their domain. After the reform, panels started competing for the overall budget regardless of the domain; all panels are then assigned to a single group regardless of their domain.

We stipulate that the evaluation noise in a given panel is constant over the period of analysis. However, the difference in the pool of competing panels over time implies that panels face a change in their relative evaluation noise. We relate applications to a given panel i to the reviewer noise in that panel as well as the noise in the panels in the same group. For the empirical analysis, we hypothesize that the number of applications to a panel depends on the difference $\sigma_i - \bar{\sigma}_t(i)$, where $\bar{\sigma}_t(i)$ is the average of the reviewer noise in the relevant group to which the panel belongs, G_{it}. The reform induced a change in the relative evaluation noise that is time- and panel-specific. We therefore estimate the following econometric model with a difference-in-difference structure where the intensity of treatment varies across panels and time periods:

$$(16) \quad a_{ist} = \alpha_{is} + \alpha_t + \beta_a[\sigma_i - \bar{\sigma}_t(i)] + \varepsilon_{ist}, \qquad \bar{\sigma}_t(i) = \frac{1}{N_{it}} \sum_{j \in G_{it}} \sigma_j.$$

We allow for panel times seniority fixed effects (α_{is}) and year fixed effects (α_t). The identifying variation derives from the reform that changed the funding allocation and its specific effect across panels. The regression assumes that the disturbances ε_{ist} are potentially heteroskedastic, contemporaneously correlated across panels, and autocorrelated. We supplement the analysis by relating the resulting share of the budget allocation to each panel and seniority call, B_{ist}, to evaluation noise in a similar way as in equation (16) and we denote the marginal effect of evaluation noise on the budget share by β_B.

V.C. Measuring Evaluation Noise: RCN Data and Matching with ERC Panels

To estimate equation (16), we construct a measure of evaluation noise σ_i for each panel. We measure the evaluation noise as the disagreement among grades of reviewers evaluating grant applications in a specific panel. Panels with a larger share of reviewers who agree on a grade are panels with lower noise.

Given that reviewer evaluation at the ERC is not available, we quantify evaluation noise by research field using data on grant evaluations at the RCN (see the tasks listed in step 1 in Figure VII). We obtained complete data, including grades for the universe of RCN applications, successful or not, over an extended period from 2002 to 2021. We focus on all the proposals submitted within the FRIPRO program, which funds curiosity-driven academic research proposed by researchers across all disciplines, similar to the ERC. The median amount of the grant is about \$1.3 million in 2020 and is awarded to individual researchers rather than research centers. The referees are mostly senior scholars of international standing, with a median age of 52 and based in 42 different countries, with the United Kingdom, Germany, and Sweden being the most frequent countries of origin in the 2020–2021 period. Given that the setting is comparable to the ERC, we assume that the evaluation dispersion at the RCN and the ERC are similar.

The next step consists in assigning each RCN application to one of the 25 ERC panels to then compute reviewer agreements in the corresponding ERC panels. We measure reviewer agreements on the basis of the grades of each application in a given panel. We explain each step below.

The FRIPRO program is divided into broad domains based on the applicant's fields. First, given that FRIPRO domains are similar to those used at the ERC, it is straightforward to assign each application to one of the ERC domains (LS, PE, and SH). Second, we assign each RCN application to one of the ERC panels (6 panels in the SH domain, 10 in PE, and 9 in LS) in the corresponding domain as follows. Exploiting text information, we construct a prediction algorithm to assign applications to panels using text from both titles and abstracts of ERC and RCN applications. We use 10,962 ERC applications (corresponding to the universe of successful applications between 2007 and 2020), and 9,964 RCN applications (including both successful and unsuccessful applications). In both sets of applications, the text (title and abstract) describing the research project has a total of about 2,100 characters, corresponding to about 300 words. As explained in detail in Online Appendix C, we perform the classification of each application using machine learning techniques combining BERT

(Bidirectional Encoder Representations from Transformers, Devlin et al. 2019) with a neural network algorithm. The validation accuracy ranges from 74% to 83%. We allocate each application to an ERC panel based on the highest predicted probability in the main analysis. As a robustness check, we perform a bootstrap analysis where we randomly assign a given application to a panel according to the vector of probabilities of belonging to a particular panel.

Having assigned RCN applications to ERC panels, we develop a measure of evaluation noise for each ERC panel based on the agreement among reviewer grades in the RCN applications assigned to that panel (our second task listed in step 1 in Figure VII). Overall, we have 40,156 observations of RCN reviewer grades, or about 4.05 grades per application.[39]

We borrow methods and statistics developed in education and psychology for measuring reviewer agreement based on comparing the grading patterns of multiple evaluators. The simplest measure is the percentage agreement between reviewers, that is, the number of times any pair of reviewers agree on the same grade divided by the number of possible pairs. This measure tends to overestimate the amount of agreement because it also includes agreement that would occur by chance (Cohen 1960, 1968). To adjust for chance agreement, measures such as Cohen's kappa, Fleiss's kappa, Gwet's AC, and Brennan's AC have been developed; see Gwet (2014) for a review and Online Appendix C for further details. We compute these different measures and find them to be highly correlated with each other, with pairwise correlations ranging between 0.73 and 0.99.

For illustration, Figure IX plots the inter-rater agreement computed as Gwet's AC. The largest agreement among reviewers is in panels in the PE domain, particularly in PE01 (mathematics) and PE09 (universe sciences). In contrast, the lowest agreement occurs in the SH domain, especially in SH05 (cultural studies) and SH06 (history). Online Appendix Table IV provides a complete list of different agreement measures across all panels, with standard errors. We use minus the inter-rater agreement measure to measure evaluation noise.

39. Online Appendix C provides further descriptive statistics on the data we use for the analysis.

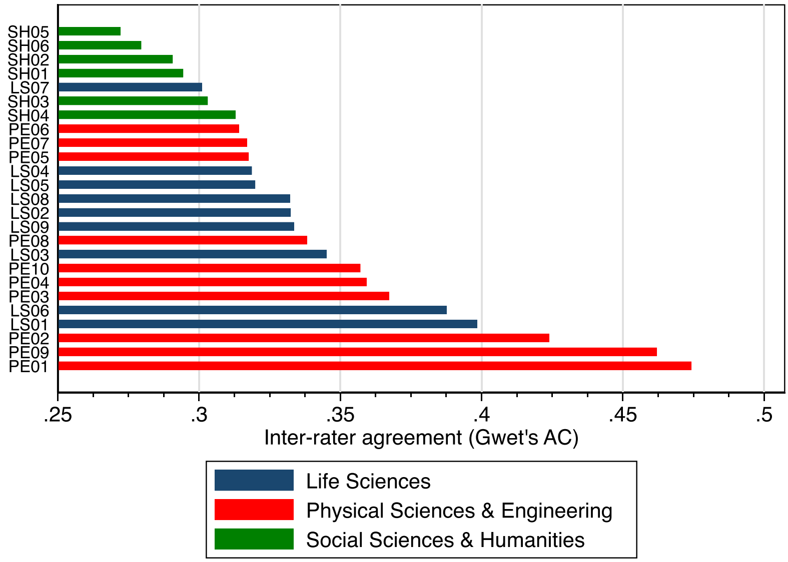

FIGURE IX

Inter-Rater Agreement by ERC Panel

This figure shows the inter-rater agreement, computed as Gwet's AC, based on reviewer grades for RCN funding applications. Each application has been assigned to an ERC panel based on text analysis. LS: Life Sciences, PE: Physical Sciences and Engineering, SH: Social Sciences and Humanities. SH05 covers cultural studies, whereas PE01 covers mathematics. *Source:* RCN and ERC data.

V.D. Effect of Evaluation Noise on Applications

The estimated coefficients β_a and β_B from equation (16) are displayed in Table I. The table displays five regressions based on different definitions of the evaluation noise. All effects are expressed in standard deviation changes. The first column displays the results for grant applications. A 1 standard deviation increase in the evaluation noise in a field increases applications in that field by 0.4 to 0.6 of a standard deviation. The effect is similar when using different definitions of reviewer agreement (and therefore evaluation noise). Given the econometric model, the comparative statics predictions of Proposition 3 translate to

$$\text{(17)} \qquad \frac{\partial a_{ist}}{\partial \sigma_i} = \beta_a > 0 \qquad \frac{\partial a_{ist}}{\partial \sigma_{i'}} = -\frac{\beta_a}{N_{it}} < 0,\ i' \neq i, i' \in G_{it}.$$

We thus find empirical confirmation for the predictions of our theory. Turning to the effect of evaluation noise on grant allocation

TABLE I
EFFECT OF EVALUATION NOISE ON FUNDING OUTCOMES

Evaluation noise based on:	Requested funding (β_a)	Budget shares (β_B)
Percent agreement	0.41***	0.428***
	(0.093)	(0.131)
Cohen's kappa	0.583***	0.462**
	(0.162)	(0.202)
Fleiss's kappa	0.577***	0.448**
	(0.163)	(0.201)
Gwet's AC	0.402***	0.425***
	(0.09)	(0.128)
Brennan's AC	0.41***	0.428***
	(0.093)	(0.131)
Observations	845	850

Notes. This table shows the effect of evaluation noise computed from different inter-rater agreement measures on the requested funding and allocated budget shares, see equation (16). The inter-rater agreement measures are computed using a power weighting scheme. The coefficients are expressed in standard deviation effects. Each cell is a separate regression, controlling for time and panel × seniority fixed effects. Panel-corrected standard errors are calculated using a Prais-Winsten regression, where a panel × seniority–specific AR(1) process is assumed. This also allows the error terms to be panel × seniority specific, heteroskedastic, and contemporaneously correlated across panels × seniority groups. Significance level: * $p < .1$, ** $p < .05$, *** $p < .01$.

shares, we also find consistent and statistically significant effects indicating that budget shares are increasing in their own evaluation noise. A standard deviation increase in own evaluation noise increases a field's budget by 0.4 to 0.5 of a standard deviation or by 0.5% from a baseline of 4%. In all four specifications, the reform led to a significant change in the level of funding (and therefore applications), with the ERC fields with the lowest evaluation noise lagging. Online Appendix Table V shows that the results are robust to using a probabilistic classification of proposals into ERC panels. It also provide an event analysis that shows the absence of pretrends.

Overall, we conclude that a simple model that accounts for differences in evaluation noise across fields can explain the changes in the ERC budget allocations that occurred after the reform, even at the finer 25-panel subdivision.

VI. ENDOGENOUS EVALUATION NOISE: GAME AMONG FIELDS

Our baseline analysis takes the evaluation noise in each field as exogenously given. However—given that under proportional apportionment the application level and thus the budget allocated to a field increases in noise—each field acting as a collective might be tempted to raise its noise level, for instance, by reducing the quality of panelists. Coordination at each field level could be achieved through a representative appointed by the scholarly association in the field. Similarly, in the grading application, teachers in each course could easily add noise to their grades. To analyze these situations, this section sketches a game-theoretic extension of the model in which fields independently choose the noise level in their evaluation process.

As a proof of concept, consider two fields, $i = 1, 2$, with additive noise and an identical distribution of types, $G_i(\theta) = G(\theta)$. In the first stage, suppose that each field i acts as a player and simultaneously sets its noise level σ_i aiming to maximize the merit of the funded projects in the field,

$$U_i\left(\sigma_i, \sigma_{-i}\right) := \int_{G^{-1}(1-a_i)}^{\overline{\theta}} \theta\left[1 - F\left(\frac{G^{-1}(1-a_i) + \sigma_i F^{-1}(1-\frac{c}{v}) - \theta}{\sigma_i}\right)\right] g(\theta) d\theta, \tag{18}$$

where $a_i = a_i(\sigma_i, \sigma_{-i})$ is the level of applications that result in the general equilibrium in the second stage and σ_{-i} is the noise in the other field.[40] Suppose that the action set for field i is $[\sigma_i^0, \infty)$: field i can voluntarily increase its level of noise, but cannot decrease it below a set level σ_i^0 corresponding to the field's initial "intrinsic" noise level. While decreasing the level of noise is prohibitively costly, the field can freely increase the level of noise above σ_i^0.

The noise levels (σ_i, σ_{-i}) chosen in the first-stage game are publicly observed. In the second stage, candidates in each field apply, and the total budget of B is allocated to the two fields in proportion to applications. For any given (σ_i, σ_{-i}), the second-stage equilibrium is then determined by the solution of the system

$$\int_{G^{-1}(1-a_i)}^{\overline{\theta}} \left[1 - F\left(\frac{G^{-1}(1-a_i) + \sigma_i F^{-1}\left(1-\frac{c}{v}\right) - \theta}{\sigma_i}\right)\right] g(\theta) d\theta = \frac{a_i}{a_1 + a_2} B$$

for $i = 1, 2$.

40. If, instead, fields only cared about maximizing the number of grants assigned to their field, each field would aim to make its signal as noisy as possible.

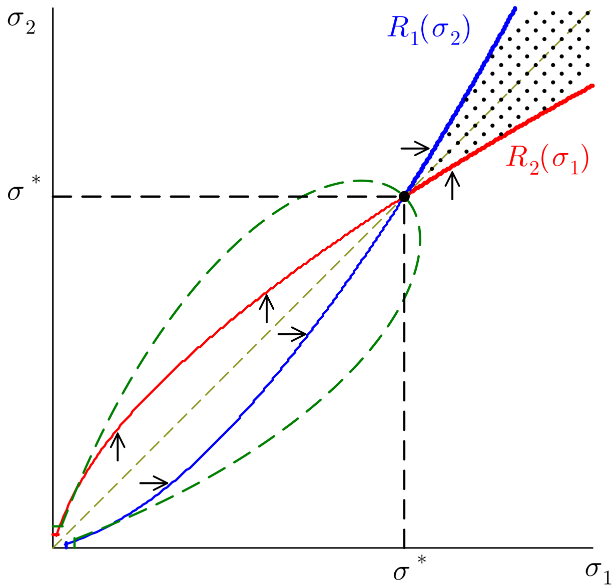

FIGURE X

Equilibrium Regimes in the Field Game

The best replies are depicted in blue and red. The level curve for the total payoff at (σ^*, σ^*) is in green.

For an example with ε_i and θ_i normally distributed, Figure X displays field 2's best reply $\sigma_2 = R_2(\sigma_1)$ in red as a function of field 1's noise and similarly $\sigma_1 = R_1(\sigma_2)$ in blue. To understand the shape of the best replies, note that when evaluation in the other field is perfect, $\sigma_{-i} = 0$, it is enough for field i to set an infinitesimal σ_i to obtain the entire budget. As σ_{-i} increases, field i obtains less budget, resulting in reduced applications. Then, provided that the expected merit of the marginally funded applicant is positive, it becomes optimal to increase the noise to obtain a larger budget. On the one hand, holding fixed the level of applications, an increase in noise reduces the effectiveness of evaluation and thus has a negative direct effect on the field's payoff—this effect becomes stronger as noise increases. On the other hand, as noise increases, equilibrium applications rise, in turn increasing the budget allocated to the field. At the best reply level of noise, the negative effect associated with the reduced quality of winning candidates is exactly offset by the positive marginal effect of obtaining more budget. Raising the level of noise past this level reduces the field's payoff. Best replies are upward-sloping for low noise levels and concave—increasing noise has diminishing marginal returns.[41]

Depending on the initial level of noise $\sigma^0 = (\sigma_1^0, \sigma_2^0)$, there is always a unique stable equilibrium, with a basin of attraction equal to the entire action profile. As illustrated in Figure X, there exists a benchmark level of noise $\sigma^* > 0$, such that there are three equilibrium regimes depending on the parameters:

i. Low initial noise in all fields: When $\sigma_1^0 \leqslant \sigma^*$ and $\sigma_2^0 \leqslant \sigma^*$, both fields sets their noise to the same level σ^*, resulting in the symmetric equilibrium (σ^*, σ^*). For these initial conditions, noise is equalized across fields in the equilibrium of the game.
ii. Highly asymmetric initial noise across fields: When $\sigma_1^0 \geqslant \sigma^*$ and $\sigma_2^0 \leqslant R_2(\sigma_1^0)$, the equilibrium is on the red curve $(\sigma_1^0, R_1(\sigma_1^0))$, as illustrated by the vertical arrows. In this case, field 2 increases its noise up to $R_2(\sigma_1^0)$, while field 1, which would prefer to decrease its noise, keeps it at the initial σ_1^0. Symmetrically, when $\sigma^* \leqslant \sigma_2^0$ and $\sigma_1^0 \leqslant R_1(\sigma_2^0)$, the equilibrium is on the blue curve $(R_1(\sigma_2^0), \sigma_2^0)$, as illustrated by the horizontal arrows. For parameters in this second region, the increase in noise by the less noisy field reduces only part of the initial asymmetry in noise. Part of the initial asymmetry persists in equilibrium.
iii. High initial noise in all fields: When $\sigma_1^0 \geqslant R_1(\sigma_2^0)$ and $\sigma_2^0 \geqslant R_1(\sigma_1^0)$ (and thus $\sigma_1^0 \geqslant \sigma^*$ and $\sigma_2^0 \geqslant \sigma^*$), both fields do not modify their noise levels. For these parameters, the equilibrium is at (σ_1^0, σ_2^0), equal to the initial level in both fields. All the initial asymmetry in noise persists.

What is the effect of the increase in noise on the total payoff in the two fields, $U_1(\sigma_1, \sigma_2) + U_2(\sigma_1, \sigma_2)$? The solid curve corresponds to the level curves of the total payoff achieved at (σ^*, σ^*). Strikingly, we conclude that when fields are only allowed to increase (but not decrease) their noise, starting from a relatively low but sufficiently asymmetric level of initial noise, the addition of noise in the field game can generate a gain in total

41. Best replies can eventually decrease if the expected type in the population of candidates is negative and the fraction of applicants is sufficiently high, for example, because the budget is high relatively to $\frac{c}{v}$.

payoff. For example, suppose that initially, the noise levels are $(\sigma_1^0, \sigma_2^0) = (1, \frac{1}{3})$ outside the green isopayoff but within the parameters that lead to (σ^*, σ^*). The social planner gains by allowing field 2 to raise optimally its level of noise to $R_2(1)$, to which field 1 replies with $R_1(R_2(1))$, eventually reaching (σ^*, σ^*) with total payoff

$$U_1(\sigma^*, \sigma^*) + U_2(\sigma^*, \sigma^*) > U_1\big(\sigma_1^0, \sigma_2^0\big) + U_2\big(\sigma_1^0, \sigma_2^0\big).$$

In this second-best world, the improvement in efficiency associated with a more balanced allocation of the budget across fields is larger than the reduction in efficiency due to the less meritocratic allocation within fields.

VII. Sorting across Fields/Courses

To isolate the effect of the supply-side interdependence induced by the budget allocation rule, our baseline model prevents candidates from choosing the field in which to apply. However, in the context of grantmaking, researchers who work at the crossroads between fields often have some leeway in choosing the field where they stand a better chance of funding. Similarly, university students, when selecting their major field and elective courses, might take into account their chance of obtaining an honors degree, which is typically awarded to the top 10% or 15% of students in the class. This section extends the model to incorporate the demand-side interdependence generated by the ability of candidates to select which field to apply in—or which course to enroll in.

In the spirit of Roy (1951), suppose that candidates are characterized by two dimensions of talent, θ_1 and θ_2, with identical and independent distributions, $G_i(\theta_i) = G(\theta_i)$. Candidates choose to apply in either of two fields, where field i evaluates dimension i of the applicant's θ through the noisy signal $x_i = \theta_i + \sigma_i\varepsilon$ satisfying the MLR property. For example, candidates who apply to physics are evaluated in terms of their mathematical talent, whereas verbal talent matters for literature candidates.

On the supply side, awards are allocated either (i) through a fixed budget, $B_i < \frac{1}{2}$, or (ii) in proportion to applications, a_iB, where $B < 1$ represents the total budget and a_i the number of applications in field i.[42] On the demand side, for simplicity we set the application cost to zero in both fields, $c_i = 0$. Nevertheless, given that candidates can submit a single application, they face an opportunity cost equal to the forgone probability of winning a grant in the other field. In equilibrium, candidates choose the field that maximizes their winning probability.

With either budget allocation rule, by the MLR property, the evaluator implements a cutoff acceptance policy to assign grants according to $x_i \geqslant \hat{x}_i$. The equilibrium is characterized by (i) the demand-side indifference condition

$$1 - F\left(\frac{\hat{x}_1 - \theta_1}{\sigma_1}\right) = 1 - F\left(\frac{\hat{x}_2 - \theta_2}{\sigma_2}\right), \tag{19}$$

which defines an upward-sloping indifference boundary $\hat{\theta}_2(\theta_1) = \hat{x}_2 + \frac{\sigma_2}{\sigma_1}(\theta_1 - \hat{x}_1)$ in the space $(\theta_1,\, \theta_2)$ such that for any given θ_1 types $\theta_2 \leqslant \hat{\theta}_2(\theta_1)$ apply to field 1 and otherwise apply to field 2 and (ii) supply-side budget equations for each field

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\theta}}^{\hat{\theta}_2(\theta_1)} \left[1 - F\left(\frac{\hat{x}_1 - \theta_1}{\sigma_1}\right)\right] g(\theta_1) g(\theta_2)\, d\theta_2 d\theta_1 = B_1,$$

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\hat{\theta}_2(\theta_1)}^{\bar{\theta}} \left[1 - F\left(\frac{\hat{x}_2 - \theta_2}{\sigma_2}\right)\right] g(\theta_1) g(\theta_2)\, d\theta_2 d\theta_1 = B_2.$$

To illustrate the construction, start from initial noise levels σ_1 and σ_2, resulting in equilibrium acceptance standards $\hat{x}_1$ and $\hat{x}_2$. The solid black line in Figure XI illustrates the indifference boundary resulting in symmetric noise $\sigma_1 = \sigma_2$ and acceptance standards $\hat{x}_1 = \hat{x}_2$, with axes expressed in terms of type percentiles $(G(\theta_1), G(\theta_2))$.[43] What is the effect of an increase in $\frac{\sigma_2}{\sigma_1}$, the noise in field 2 relative to field 1, on equilibrium applications in the two fields for the case with a fixed budget?

First, the change in relative noise has an effect on selection. Holding fixed the acceptance standards $(\hat{x}_1, \hat{x}_2)$, the increase in $\frac{\sigma_2}{\sigma_1}$ induces an anticlockwise rotation of the indifference boundary equation (19) around $(G(\hat{x}_1), G(\hat{x}_2))$, corresponding to the dotted green curve. Types in the upper right region to the right of the dashed red curve and above the black 45-degree line—which

42. If the budget were abundant $B > 1$, grants would always be awarded to the entire population.

43. Beyond the symmetric case, the indifference boundary $\hat{\theta}_2(\theta_1)$ in the type percentile space is nonlinear.

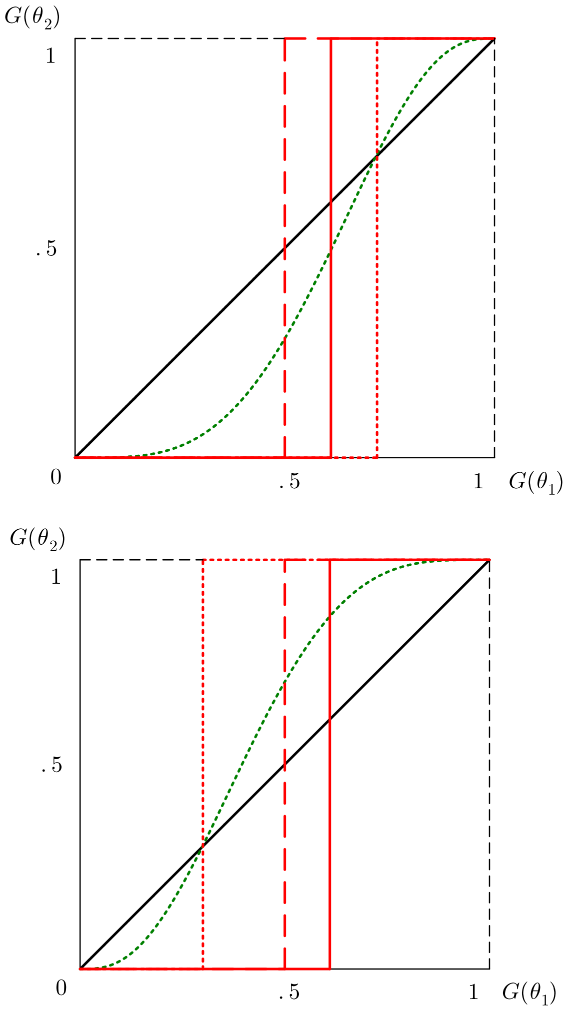

FIGURE XI

Comparative Statics with Respect to an Increase in Noise σ_2

The top panel represents case (a) in which the increase in noise in step 1 results in a reduction in a_2 holding fixed $(\hat{x}_1, \hat{x}_2)$ at the initial level. The bottom panel represents the opposite case.

are highly talented in both dimensions—have the incentive to flee the relatively noisier field 2 and join field 1, where they are now relatively more likely to clear the acceptance bar. Intuitively, the winning probability of these candidates is now higher in the relatively more meritocratic field 1, even though they are even more talented in dimension θ_2 than θ_1. At the same time, candidates in the lower left region (to the left of the dashed red curve and below the black 45-degree line) with lower talent in dimension θ_1 now find the noisier field 2 more attractive, even though they are relatively worse in dimension θ_2 than θ_1. Overall, the more meritocratic field 1 attracts more talented candidates, and less talented candidates prefer to hide in the noisier field 2.

To understand how noise affects the level of applications in the two fields, note that as a result of the first step, application levels either (a) decrease or (b) increase, depending on the relative size of the regions of types switching field, as represented respectively by Figure XI. As a proof of concept, consider the extreme case in which evaluation becomes perfect in field 1, $\sigma_1' = 0$, resulting in a vertical indifference boundary (red dotted curves). The second step consists in adjusting the acceptance standard in field 1 until applications in field 1 are reset to the initial level. This is achieved at $\hat{x}_1 = G^{-1}(\frac{1}{2})$, given that we started from a symmetric situation. In case (a), $\hat{x}_1$ should be reduced to increase applications by translating the indifference boundary to the dashed red line—by construction the area to the right of the dashed curve and to the left of the black 45-degree line (high-merit applications gained) is equal to the area to the left of the dashed curve and to the right of the black 45-degree line (low-merit applications lost). A similar construction applies to case (b), when $\hat{x}_1$ should instead be increased to move the indifference boundary to the left and thus reduce applications in field 1.

Third, with perfect information, all applicants in field 1 are awarded a grant for sure.[44] Having restored applications to the initial level, grant awards would be $\frac{1}{2} > B_1$, thus overspending the initial budget. When the budget is fixed, to re-equilibrate the imbalance in the budget, $\hat{x}_1$ must necessarily increase relative to the level in the second step, shifting the indifference boundary to

44. In general, the composition of applicants in field 1 has now improved in the first-order stochastic order. The density for types below (above) a critical level $\tilde{\theta}$ is reduced (increased), implying stochastic dominance.

the right, as represented by the red curves in the two panels of figures. Hence, in the new equilibrium a_1 must decrease to $a_1' = B_1 < \frac{1}{2}$ and thus a_2 must increase to $a_1' = 1 - B_1 > \frac{1}{2}$.

Finally, turn to the outcome under proportional budget allocation. Unraveling results in field i when evaluation is perfect in that field ($\sigma_i = 0$) or completely noisy in the other field ($\sigma_{-i} \to \infty$). Under proportional apportionment, all types can guarantee the average winning probability by applying to field 2—thus, candidates who would win with a probability below this level leave field 1, and the process continues until field 1 unravels, $a_1 = 0$. More generally, when candidates choose their field with normal types and normal noise we verified numerically that equilibrium applications increase in the field's noise and decrease in the noise in the other field, and the effect is stronger under a proportional than under a fixed budget.[45]

VIII. Organization of Funding with Noisy Evaluation

VIII.A. Design of Funding Rules

The optimal allocation for the grantmaker maximizes the total merit across fields

$$\sum_{i=1}^{N} \int_{G_i^{-1}(1-a_i)}^{\bar{\theta}_i} \theta \left[1 - F_i\left(\frac{\hat{x}^D(a_i) - \theta}{\sigma_i}\right)\right] g_i(\theta)\, d\theta,$$

subject to the demand system $\hat{x}_i^D(a_i) = G_i^{-1}(1 - a_i) + \sigma_i F_i^{-1}(1 - \frac{c_i}{v_i})$ given the total budget available for distribution

$$\sum_{i=1}^{N} A_i(a_i) = \sum_{i=1}^{N} \int_{G_i^{-1}(1-a)}^{\bar{\theta}_i} \left[1 - F_i\left(\frac{\hat{x}_i^D(a_i) - \theta}{\sigma}\right)\right] g_i(\theta)\, d\theta = B.$$

To illustrate how the equilibrium compares to the optimal allocation, consider initially two symmetric fields with normally distributed types and signals and a PA budget rule. The identical

45. We verified that the main results of the article extend to the field choice model with normal noise when types in each field follow the generalized normal distribution (also known as exponential power) with density $g(\theta) = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} e^{-(\frac{|\theta-\mu|}{\alpha})^{\beta}}$, encompassing the Laplace ($\beta = 1$), normal ($\beta = 2$), and uniform ($\beta \to \infty$) distributions. When the type distribution has an increasing hazard rate ($\beta > 1$), the equilibrium is unique; unraveling results when the upper tail is thicker than exponential, $\beta < 1$.

equilibria in the first and second field is represented in Figure XII by the blue dot marked as (i) at the crossing of the solid red curve $A_i(a_i)$ and the dotted blue curve $B_i(a_i)$. In this symmetric setting, the PA equilibrium allocation is also optimal.

As noise dispersion σ_2 in the second field increases, the award function in field 2 shifts down to the dashed red curve in the bottom panel, so that the partial equilibrium applications increase in field 2 from the initial level corresponding to the blue dot (i) to the green dot. The top panel shows the reduction in the budget in field 1 to the dot-dashed blue curve due to the increase in applications in field 2, as we adjust to the general equilibrium represented by the red dot (ii). In turn, this reduction in applications in field 1 increases the budget available in field 2 to the dot-dashed blue curve in the bottom panel, leading to a further increase in applications, eventually resulting in a new general equilibrium at the red dot (ii).

As noise increases in field 2, it becomes optimal for the grantmaker to transfer some of the overall budget from the noisier field 2 to the relatively more accurate field 1, resulting in the grantmaker optimal allocation (iii) marked by the black dots. Departing from proportional allocation PA, the grantmaker can implement this optimal allocation within the QPA class of budget rules by reducing proportionality ϱ_1 in the first field.

VIII.B. Pooling Fields

In the baseline model, each panel evaluates a single field and is characterized by a field-specific level of evaluation noise for all applicants in the panel. In reality, panels at research funding organizations are typically assigned applications that belong to different fields. What is the effect of pooling heterogeneous fields into a single panel relative to assigning each field to a separate panel?

As a proof of concept, suppose there are two fields in the same discipline. Think of the basic and clinical research in a medical specialty, such as pancreatic cancer research. Suppose that evaluation is noisier for clinical than for basic research, $\sigma_2 > \sigma_1$.

It is useful to reinterpret the selection of grantees in a constant payline equilibrium for a single field as follows. Express the acceptance standard, rather than in terms of the signal x, in terms of the corresponding posterior expectation $E_{\sigma_i}[\theta|a_i, x]$ about the application's merit θ computed via Bayes's rule. Given a_i, the

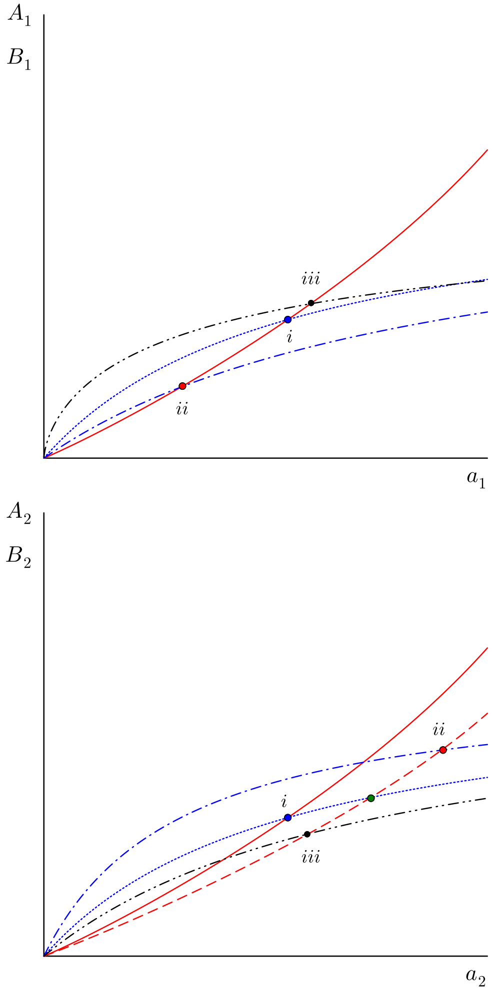

FIGURE XII

Design of Responsiveness of Allocation Rule

The top and bottom panels represent equilibria in fields 1 and 2, respectively, where award and budget functions cross. The blue dots (i) are the initial symmetric equilibria. The red dots (ii) are the equilibria with PA allocation following an increase in noise in field 2. The black dots (iii) are the optimal allocations, which can be implemented with a subproportional budget rule in which the budget is less responsive in field 2 than 1.

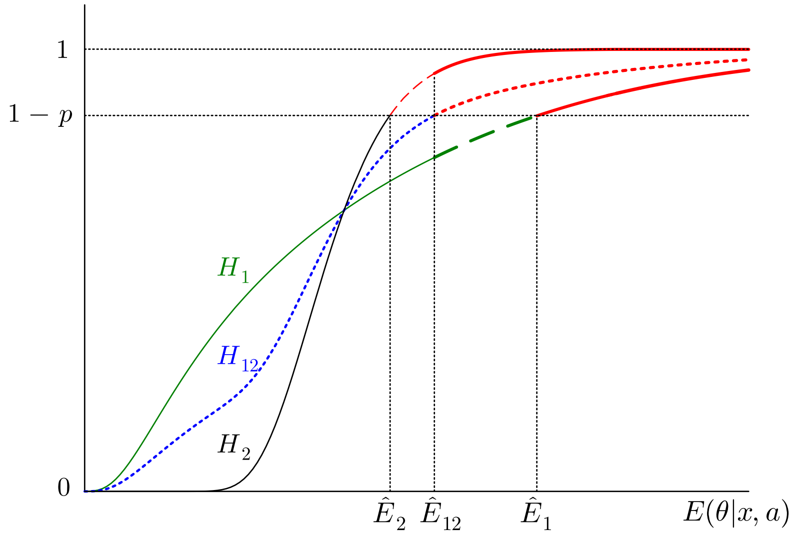

FIGURE XIII

Effect of Pooling Fields

The distributions of scores for fields 1 and 2 are represented in green (H_1) and black (H_2), respectively, When fields are evaluated in isolation, grantees correspond to the top segments of the separate distributions of scores for fields 1 (green) and 2 (black). When fields are merged, scores follow the mixture distribution (dotted blue) resulting in a loss of awards in field 1 (thin dashed red) and a gain in awards in field 2 (thick dashed green).

constant payline acceptance standard expressed in terms of the posterior expectation (or score) $E_{\sigma_i}[\theta|a_i, x]$ is then

$$1 - H_i(\hat{E}_i) = p.$$

Given that the score $E_{\sigma_i}[\theta|a_i, x]$ is an increasing function of x by the MLR property, the two representations are equivalent when all applicants are evaluated with a common signal structure F_{σ_i}, as in the baseline model. The expected merit score of the marginally accepted candidate satisfies $E_{\sigma_i}[\theta|a_i, \hat{x}_i] = \hat{E}_i$, linking $\hat{x}_i$ and $\hat{E}_i$.

For a field evaluated in isolation with noise σ_1 and given application level a_1, the score $E_{\sigma_1}[\theta|a_1, x]$ is distributed according to H_1. Under constant payline the marginal score is $\hat{E}_1$, as illustrated by the green curve H_1 in Figure XIII when both types and signals are normally distributed. Similarly, for a noisier field evaluated in isolation, the black curve corresponds to distribution H_2 of scores $E_{\sigma_2}[\theta|a_2, x]$, resulting in marginal score $\hat{E}_2$. As illustrated for this example, a reduction in noise (or increase in accuracy) induces a mean-preserving clockwise rotation in the score distribution.[46]

Turn now to the case in which applicants in the two fields are pooled in the same panel. Suppose that applications are still evaluated by the same experts in each field and that σ_1 and σ_2 remain unaffected. The pooled scores in the joint panel are distributed according to a mixture of H_1 and H_2, with weights determined by the relative level of applications in the two fields

$$H_{12} = \frac{a_1}{a_1 + a_2} H_1 + \frac{a_2}{a_1 + a_2} H_2,$$

corresponding to the dotted blue curve H_{12} in Figure XIII for the case $a_1 = a_2$. The resulting marginal score for a given payline p is now $\hat{E}_{12}$, solving $1 - H_{12}(\hat{E}_{12}) = p$.

For a realistically low payline—when the winning scores with pooled fields are above the rotation point—we have

$$\hat{E}_2 < \hat{E}_{12} < \hat{E}_1,$$

as illustrated in Figure XIII. Intuitively, the winning proposals above the payline disproportionately originate from the more accurate field a, where scores are more extreme. Applicants in field 1 with scores between $\hat{E}_{12}$ and $\hat{E}_1$ (the green dashed segment of H_1 in Figure XIII) are now awarded grants at the expense of applicants in field 2 with scores between $\hat{E}_2$ and $\hat{E}_{12}$ (red dashed segment of H_2). The more accurate field can increase the fraction of successful applications above the payline p and thus enjoys a higher effective payline, $1 - H_1(\hat{E}_{12})$. Conversely, the noisier field experiences a lower effective payline, $1 - H_2(\hat{E}_{12})$.

Through this mechanism, pooling fields with heterogeneous noise dampens the perverse effect of meritocracy on the level of applications. The more consensual field obtains the lion's share of grants in the panel. This pattern is in line with Martin, Lindquist, and Kotchen's (2008) empirical finding that basic research has a higher success rate than clinical research at the NIH, where paylines across panels (also known as study sections) are nevertheless equalized. Clinical research suffers from being less consen-

46. In the limit as signal noise $\sigma \to \infty$, the distribution of the posterior expectation becomes a step function at the prior $E[\theta \,|\, a]$. As $\sigma \to 0$, the distribution of the posterior expectation converges to the prior distribution, $G(\theta \,|\, a)$.

sual because it is pooled with basic research in the same panels, consistent with our prediction. If clinical studies and basic science were regrouped in separate panels, their success rate would automatically equalize. However, according to our analysis, more applications would be submitted for clinical studies and fewer for basic science.

Noisier fields thus have a strong incentive to split from more consensual fields and lobby to have their separate panel. As a result, not only will the fraction of accepted applications increase for noisier fields that set up their panel, incentives to apply will also be stepped up, resulting in an increase in awards. Conversely, more consensual fields prefer to be merged with noisier fields.[47]

Under proportional allocation, fields that are assigned to separate panels have a perverse incentive to increase noise relative to the other panels, whereas pooling with other fields induces a virtuous incentive to decrease their noise relative to other fields in the same panel, thus gaining awards at the expense of other fields in the same panel.

VIII.C. *Benchmarking*

This logic can also shed light on a benchmarking practice adopted by the NIH, according to which percentiles are computed by pooling scores across recent evaluation cycles at the same panel, also known as study sections. As explained by the National Institute of Health (1988), percentiles for applications in each evaluation cycle are calculated by pooling current scores with scores given by the same study sections to the applications evaluated in the preceding two cycles, a system that is still in operation today.[48]

What might look like an inessential tweak to the payline system has important consequences. If some applications were submitted in the previous two cycles, $a_{t-1} + a_{t-2} > 0$, even if in the current cycle evaluation were perfect, $\sigma_t = 0$, some budget would always be available for distribution. Hence, unraveling would not result. More generally, benchmarking dampens the effect of noise

47. Clearly, the benefit of pooling for the more consensual fields could be dampened if applications were assigned to less accurate reviewers and the accuracy of evaluation were to decrease.

48. See https://www.niaid.nih.gov/grants-contracts/understand-paylines-percentiles for a detailed account.

on applications by reducing the responsiveness of the budget to applications.

Similar to pooling, benchmarking can reverse the perverse comparative statics of proportional allocation with respect to noise. By improving its accuracy in this cycle compared to the previous cycle, a panel can increase the fraction of successful applications above the payline. Under the reasonable assumption that reviewers aim to assign as many grants as possible to applicants in their panel (possibly at the expense of other panels), they now have the incentive to be more accurate than in the previous cycle so as to increase dispersion in the posterior expectation and thus increase the number of funded applications in the panel. Through this channel, the NIH method of computing percentiles relative to the applications previously evaluated by the same panel incentivizes accurate evaluation, triggering virtuous incentives to increase accuracy, in contrast to the vicious incentives highlighted in our baseline analysis.

IX. Contribution to Literature

Grantmaking has received relatively limited attention from economists. While our analysis predominantly takes a positive approach to commonly employed nonmarket resource allocation methods, previous research has primarily concentrated on normative considerations. In a pioneering application of marginal analysis, Peirce (1879) sketches the normative theory of resource allocation across research fields for a planner. As stressed at least since Arrow (1962), market forces tend to underprovide research, mostly because invention is nonrival. Governments, however, have limited information about the benefits of research in different fields. Weisbrod (1961) offers an early attempt to evaluate the social benefits of medical research across diseases.[49] Weinstein and Zeckhauser (1973) link the problem of the optimal allocation of budget to fields to the decision-theoretic approach underlying hypothesis testing.

Turning to positive analyses, Wildavsky (1964) describes the incremental nature of the budget apportionment process for determining government funding of the NIH in the early

49. In a review of the NIH, Zeckhauser (1967) also argues that disease burden should guide funding choices.

days; our static model abstracts from dynamic considerations.[50] Zuckerman and Merton (1971) notice that acceptance rates at leading scholarly journals vary across academic disciplines, with higher rejection rates in social sciences and humanities compared with physical sciences; our analysis shows that the performance of allocation rules with proportional elements is particularly problematic when fields are heterogeneous.[51] Rejection rates also vary along similar lines across directorates at the National Science Foundation.[52]

In terms of theory, Lazear (1997) outlines a lottery model of research funding (researchers can increase their chance of obtaining a grant by buying more tickets) but abstracts away from self-selection and noisy evaluation on which we focus. Scotchmer (2004, chap. 8) formulates a simple dynamic model of demand for funding where high-type researchers self-select into applying and are disciplined to deliver because they expect to be funded in the future. Building on a setting with continuous types and scale-location signals similar to ours, Leslie (2005) sketches the demand side for submissions to academic journals—in addition to a complete analysis of the demand side, we add a (noisy) evaluation on the supply side and characterize the equilibrium depending on the budget allocation rule.[53] See also Stephan's (2012, chap. 6) discussion of science funding and Azoulay and Li's (2022) overview of the fledgling empirical literature on grant funding for science.[54]

The application cost, akin to what Nichols and Zeckhauser (1982) call an ordeal, in our model induces more worthy

50. See also the formalization by Davis, Dempster, and Wildavsky (1964). Savage (1999) gives a historical account of the influence process behind university earmarks in comparison to merit-based public funding of research.

51. Zuckerman and Merton (1971, 77) write: "the more humanistically oriented the journal, the higher the rate of rejecting manuscripts for publication; the more experimentally and observationally oriented, with an emphasis on rigor of observation and analysis, the lower the rate of rejection."

52. Cole and Cole's (1981) landmark study documents differences in interrater agreement among reviewers across fields at the NSF.

53. See also Cotton (2013) and Taylor and Yildirim (2011), focusing on discrimination issues, which we skirt.

54. Gans and Murray (2012) overview the main funding sources available for scientists (government, private firms' internal R&D, and foundations), with a focus on comparing their different disclosure and openness requirements. Boudreau et al. (2016) investigate the role of the intellectual distance between evaluators' expertise and the research proposals in systematically shaping funding outcomes.

applicants to self-select. The evaluator uses an additional noisy signal about the applicant's type, so the application cost acts as an endogenous screening device. The noise in the evaluation process thus plays a crucial role in our model as in the literature on statistical discrimination, pioneered by Phelps (1972) and surveyed by Fang and Moro (2011). In that strand, Cornell and Welch (1996) argue that competition for ranking in a tournament discriminates against candidates whom the evaluator is less informed about. Our base model moots this channel by focusing on an evaluator who is equally informed about applicants in the same field. According to our new effect, competition in a field with more noisy evaluation becomes closer to a lottery and thus encourages more applications. In turn, when the budget of grants available to a field increases in applications, the evaluator ends up inefficiently discriminating against candidates evaluated with less noise—the opposite of Cornell and Welch's outcome.

In the agency literature, Che, Dessein, and Kartik (2013), Alonso (2018), and Frankel (2021) largely focus on how to constrain optimally biased evaluators—in our model, instead, evaluators in each field are unbiased. While our model zooms in on the noisy evaluation process of applicants, the literature on tournaments and contests—from Lazear and Rosen (1981) to O'Keeffe, Viscusi, and Zeckhauser (1984), Moldovanu and Sela (2001), Che and Gale (2003), Siegel (2009), Gross and Bergstrom (2019), and Fang, Noe, and Strack (2020)—mostly focus on the incentives of contestants to exert effort, from which we abstract. Closer to our setting, Morgan, Sisak, and Várdy (2018) analyze the incentives of applicants to select different fields in a setting with exogenous supply, whereas we focus on endogenously determining the supply through the budget allocation.[55]

At a technical level, we leverage Lehmann's (1988) quantile function approach to derive sharp predictions on the effect of evaluation noise.[56] Exploiting the structure of the problem, where evaluation noise in a field affects the other fields only through the

55. We also abstract away from dynamic considerations. See Board, Meyer-ter-Vehn, and Sadzik (2020) for a model of recruitment where the accuracy of evaluation endogenously depends on past recruits; Moisson and Tirole (2020) for a foray into the dynamics of cooptation; and Bardhi, Guo, and Strulovici (2020) for characterization of when costly experimentation amplifies or dampens small differences in ability.

56. This approach is little known in economics, with the notable exception of Persico (2000).

budget allocation rule, we can obtain unambiguous comparative statics. Our results linking comparative statics to stability are in line with Samuelson's (1947) correspondence principle; see Hale et al. (2014) for an overview of the tools. Relative to the literature on fair division of resources among claimants, recently summarized by Thomson (2019), our model endogenizes the claims (applications are costly) and introduces imperfect verification (evaluation is noisy).

X. Conclusion

Our analysis emphasizes the central role of evaluation noise in nonmarket allocation processes. By developing a nonparametric approach to information, we derive the testable comparative statics prediction that applications increase in noise in all stable equilibria. In addition to empirically validating this result, we extend the analysis to allow candidates to choose a field or course, as is most relevant in applications to course selection. Noisier fields are more attractive for weaker candidates who win with lower probability, thus reinforcing our baseline comparative statics.

We also show that incentives of fields to add noise in their evaluation tend to rebalance initial asymmetries to the point of even increasing allocation efficiency in the spirit of the second best. However, when the initial noise is sufficiently high, initial asymmetries persist, as in the baseline analysis. Therefore, to maximize efficiency, budget rules should be optimally designed by making the budget allocation less responsive to applications in less noisy fields. Finally, the detrimental effect of noise on selection can be dampened by pooling fields with heterogeneous noise. When pooled with noisier fields, less noisy fields obtain the lion's share of grants because their informative scores tend to be more extreme and thus end up at the top of the score distribution.

Back to the specific PA rule that motivated our analysis, this rule appears to be fair in treating all fields in the same way by automatically equalizing the fraction of successful projects over applications across different fields. Proportional allocation also eliminates administrative discretion and political meddling in funding allocation, given that the budget allocation is determined automatically only based on relative demand from applications across fields. As another important virtue, the proportional allocation scheme has the merit of flexibly responding to demand-side signals. Despite its simplicity, we argue that formula-based funding in the general SPA class has important pitfalls when fields are heterogeneous, as they typically are.

Our analysis of proportional allocation immediately applies also to large research fellowships programs, such as the EU-wide Marie Skłodowska-Curie Action (MSCA) scheme that assigns its total budget (€6.16 billion for 2014–2020) in proportion to applications across all disciplines.[57] The drawbacks our analysis highlights are particularly severe for mechanisms that link the budget across very heterogeneous fields, as is the case for the ERC and MSCA, but perhaps less problematic for funders (like the NIH) that focus on research in a single domain (like medicine, even though NIH study sections cover a wide variety of disciplines, methodologies, and topics).[58]

The bottom-up formula-based approach to funding apportionment analyzed here can be contrasted to alternative top-down approaches, such as those prevailing at the NSF, in the United Kingdom, and Australia, where legislators discretionally allocate the budget across programs, following a yearly consultation process and a detailed proposal by the directors of the research funding organizations. Even at agencies that adopt proportional allocation, success rates for different programs and across fields are regularly published and closely monitored. While differences in success rates across fields in nonproportional systems persist over time, there is an implicit pressure to reduce the budget for fields with higher success rates in favor of fields with lower success rates.

General interest academic journals are subject to a similar pressure to allocate space to different subfields in proportion to submissions. When coeditors are given a common target acceptance rate, fields with less accurate (or consensual) evaluation will attract more submissions.[59] Similarly, university admission

57. The Canadian SSHRC Doctoral Fellowships program (covering all humanities and social sciences) also follows PA.

58. While the great majority of NIH institutes/centers adopt the payline system and publish paylines, it is only understandable that some institutes/centers at the NIH prefer not to publish their paylines, thus retaining some flexibility when treating proposals from different panels.

59. See also Akerlof's (2020) discussion of how a bias toward "hardness" can arise in science. Our analysis suggests a mechanism through which hardness prevails in a discipline, even though it is detrimental to the competition across disciplines. In our model, individual disciplines tend to be dominated by harder subfields and investigations with more accurate evaluations. When elements of

boards are tempted to admit students to different programs in proportion to applications—or to increase slots available in areas that attract more applications. Giving in to this temptation may spark a race to the bottom regarding the quality of admitted students.

BOCCONI UNIVERSITY, ITALY
BOCCONI UNIVERSITY, ITALY

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online.

DATA AVAILABILITY

The code underlying this article is available in the Harvard Dataverse, https://doi.org/10.7910/DVN/NAGMJQ (Adda and Ottaviani 2023).

REFERENCES

Achen, Alexandra C., and Paul N. Courant, "What Are Grades Made Of?," *Journal of Economic Perspectives*, 23 (2009), 77–92. https://doi.org/10.1257/jep.23.3.77

Adda, Jérôme, and Marco Ottaviani, "Replication Data for: 'Grantmaking, Grading on a Curve, and the Paradox of Relative Evaluation in Nonmarkets'," (2023), Harvard Dataverse, https://doi.org/10.7910/DVN/NAGMJQ

Akerlof, George A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84 (1970), 488–500. https://doi.org/10.2307/1879431

———, "Sins of Omission and the Practice of Economics," *Journal of Economic Literature*, 58 (2020), 405–418. https://doi.org/10.1257/jel.20191573

Alonso, Ricardo, "Recruiting and Selecting for Fit," London School of Economics Working Paper, 2018. https://doi.org/10.2139/ssrn.3124315

Arrow, Kenneth J., "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Richard R. Nelson, ed. (Princeton, NJ: Princeton University Press, 1962), 609–626. https://doi.org/10.1515/9781400879762-024

Azoulay, Pierre, and Danielle Li, "Scientific Grant Funding," in *Innovation and Public Policy*, Austan Goolsbee and Benjamin F. Jones, eds. (Chicago: University of Chicago Press, 2022), 117–150. https://doi.org/10.7208/chicago/9780226805597.003.0005

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat, "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules," *Review of Economic Studies*, 86 (2019), 117–152. https://doi.org/10.1093/restud/rdy034

proportionality are present in the allocation of resources across disciplines, disciplines with more accurate evaluation are destined to obtain fewer resources and thus become less attractive.

Bagnoli, Mark, and Ted Bergstrom, “Log-Concave Probability and Its Applications,” *Economic Theory*, 26 (2005), 445–469. https://doi.org/10.1007/s00199-004-0514-4

Bardhi, Arjada, Yingni Guo, and Bruno Strulovici, “Early-Career Discrimination: Spiraling or Self-Correcting?,” Duke University and Northwestern University Working Paper, 2020.

Biagioli, Mario, “Galileo’s System of Patronage,” *History of Science*, 28 (1990), 1–79. https://doi.org/10.1177/007327539002800101

Blackwell, David, “Comparison of Experiments,” *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1 (1951), 93–102.

Board, Simon, Moritz Meyer-ter-Vehn, and Tomasz Sadzik, “Recruiting Talent,” UCLA Working Paper, 2020.

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, and Christoph Riedl, “Looking across and Looking beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science,” *Management Science*, 62 (2016), 2765–2783. https://doi.org/10.1287/mnsc.2015.2285

Bush, Vannevar, *Science: The Endless Frontier* (Washington, DC: U.S. Government Printing Office, 1945).

Caves, Richard, *Creative Industries: Contracts between Art and Commerce* (Cambridge, MA: Harvard University Press, 2000).

Che, Yeon-Koo, and Ian Gale, “Optimal Design of Research Contests,” *American Economic Review*, 93 (2003), 646–671. https://doi.org/10.1257/000282803322157025

Che, Yeon-Koo, Wouter Dessein, and Navin Kartik, “Pandering to Persuade,” *American Economic Review*, 103 (2013), 47–79. https://doi.org/10.1257/aer.103.1.47

Cohen, Jacob, “A Coefficient of Agreement for Nominal Scales,” *Educational and Psychological Measurement*, 20 (1960), 37–46. https://doi.org/10.1177/001316446002000104

———, “Weighted Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit,” *Psychological Bulletin*, 70 (1968), 213–220. https://doi.org/10.1037/h0026256

Cole, Jonathan, and Stephen Cole, *Peer Review in the National Science Foundation: Phase Two of a Study* (Washington, DC: National Academies Press, 1981).

Cornell, Bradford, and Ivo Welch, “Culture, Information, and Screening Discrimination,” *Journal of Political Economy*, 104 (1996), 542–571. https://doi.org/10.1086/262033

Cotton, Christopher, “Submission Fees and Response Times in Academic Publishing,” *American Economic Review*, 103 (2013), 501–509. https://doi.org/10.1257/aer.103.1.501

Cullen, Julie Berry, Mark C. Long, and Randall Reback, “Jockeying for Position: Strategic High School Choice under Texas’ Top Ten Percent Plan,” *Journal of Public Economics*, 97 (2013), 32–48. https://doi.org/10.1016/j.jpubeco.2012.08.012

Crosland, Maurice, and Antonio Gálvez, “The Emergence of Research Grants within the Prize System of the French Academy of Sciences, 1795–1914,” *Social Studies of Science*, 19 (1989), 71–100. https://doi.org/10.1177/030631289019001002

Davis, Otto A., Michael A. H. Dempster, and Aaron Wildavsky, “A Theory of the Budgetary Process,” *American Political Science Review*, 60 (1966), 529–547. https://doi.org/10.2307/1952969

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, *Volume 1 (Long and Short Papers)* (Minneapolis, MN: Association for Computational Linguistics, 2019), 4171–4186.

European Commission, *ERC Work Programme 2008* (European Commission C, 2007), 5746.
Fang, Dawei, Thomas Noe, and Philipp Strack, "Turning Up the Heat: The Discouraging Effect of Competition in Contests," *Journal of Political Economy*, 128 (2020), 1940–1975. https://doi.org/10.1086/705670
Fang, Hanming, and Andrea Moro, "Theories of Statistical Discrimination and Affirmative Action: A Survey," in *Handbook of Social Economics*, Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, eds. (Amsterdam: Elsevier, 2011), 133–200. https://doi.org/10.1016/B978-0-444-53187-2.00005-X
Frankel, Alexander, "Selecting Applicants," *Econometrica*, 89 (2021), 615–645. https://doi.org/10.3982/ECTA15510
Gans, Joshua S., and Fiona E. Murray, "Funding Scientific Knowledge: Selection Disclosure and the Public-Private Portfolio," in *The Rate and Direction of Inventive Activity Revisited*, Joshua Lerner and Scott Stern, eds. (National Bureau of Economic Research, Cambridge, MA, 2012), 51–106. https://doi.org/10.7208/chicago/9780226473062.003.0005
Gross, Kevin, and Carl T. Bergstrom, "Contest Models Highlight Inherent Inefficiencies of Scientific Funding Competitions," *PLoS Biology*, 17 (2019), e3000065. https://doi.org/10.1371/journal.pbio.3000065
Gwet, Kilem Li, *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters* (Piedmont, CA: Advanced Analytics, 2014).
Hale, Douglas, George Lady, John Maybee, and James Quirk, *Nonparametric Comparative Statics and Stability* (Princeton, NJ: Princeton University Press, 2014).
Johnson, Valen E., *Grade Inflation: A Crisis in College Education*. (New York: Springer, 2013).
Karlin, Samuel, and Herman Rubin, "The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio," *Annals of Mathematical Statistics*, 27 (1956), 272–299. https://doi.org/10.1214/aoms/1177728259
Keynes, J. Maynard, "The Arts Council: Its Policy and Hopes," *Listener*, 34 (1945), 31.
Koppenol, Willem H., Patricia L. Bounds, and Chi V. Dang, "Otto Warburg's Contributions to Current Concepts of Cancer Metabolism," *Nature Reviews Cancer*, 11 (2011), 325–337. https://doi.org/10.1038/nrc3038
Lazear, Edward P., "Incentives in Basic Research," *Journal of Labor Economics*, 15 (1997), 167–197. https://doi.org/10.1086/209860
Lazear, Edward P., and Sherwin Rosen, "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89 (1981), 841–864. https://doi.org/10.1086/261010
Leat, Diana, *Philanthropic Foundations, Public Good and Public Policy* (London: Palgrave Macmillan, 2016).
Lehmann, Erich L., "Comparing Location Experiments," *Annals of Statistics*, 16 (1988), 521–533. https://doi.org/10.1214/aos/1176350818
Leslie, Derek, "Are Delays in Academic Publishing Necessary?," *American Economic Review*, 95 (2005), 407–413. https://doi.org/10.1257/0002828053828608
MacLeod, Roy M., "The Royal Society and the Government Grant: Notes on the Administration of Scientific Research, 1849–1914," *Historical Journal*, 14 (1971), 323–358. https://doi.org/10.1017/S0018246X00009638
Mandel, Richard, *Half a Century of Peer Review, 1946–1996* (Washington, DC: National Institutes of Health, 1996).
Marshall, Albert W., and Ingram Olkin, *Life Distributions* (New York: Springer Science, 2007).
Martin, Michael R., Teresa Lindquist, and Theodore A. Kotchen, "Why Are Peer Review Outcomes Less Favorable for Clinical Science than for Basic Science Grant Applications?," *American Journal of Medicine*, 121 (2008), 637–641. https://doi.org/10.1016/j.amjmed.2008.03.031
Milgrom, Paul, and Chris Shannon, "Monotone Comparative Statics," *Econometrica*, 62 (1994), 157–180. https://doi.org/10.2307/2951479

Moisson, Paul-Henri, and Jean Tirole, “Cooptation: Meritocracy vs. Homophily in Organizations,” University of Toulouse Working Paper, 2020.
Moldovanu, Benny, and Aner Sela, “The Optimal Allocation of Prizes in Contests,” *American Economic Review*, 91 (2001), 542–558. https://doi.org/10.1257/aer.91.3.542
Morgan, John, Dana Sisak, and Felix Várdy, “The Ponds Dilemma,” *Economic Journal*, 128 (2018), 1634–1682. https://doi.org/10.1111/ecoj.12473
Murphy, Richard, and Felix Weinhardt, “Top of the Class: The Importance of Ordinal Rank,” *Review of Economic Studies*, 87 (2020), 2777–2826. https://doi.org/10.1093/restud/rdaa020
National Institutes of Health, *NIH Guide for Grants and Contracts 1988. Notice: New Basis for Priority Score Percentiles at NIH*. http://grants.nih.gov/grants/guide/historical/1988_08_12_Vol_17_No_26.pdf.
Nichols, Albert L., and Richard J. Zeckhauser, “Targeting Transfers through Restrictions on Recipients,” *American Economic Review*, 72 (1982), 372–377.
O’Keeffe, Mary, W. Kip Viscusi, and Richard J. Zeckhauser, “Economic Contests: Comparative Reward Schemes,” *Journal of Labor Economics*, 2 (1984), 27–56. https://doi.org/10.1086/298022
Park, Hyunwoo, Jeongsik Lee, and Byung-Cheol Kim, “Project Selection in NIH: A Natural Experiment from ARRA,” *Research Policy*, 44 (2015), 1145–1159. https://doi.org/10.1016/j.respol.2015.03.004
Peirce, Charles S., “Note on the Theory of the Economy of Research,” *United States Coast Survey* (1879); reprinted in *Operations Research*, 15 (1967), 643–648. https://doi.org/10.1287/opre.15.4.643
Persico, Nicola, “Information Acquisition in Auctions,” *Econometrica*, 68 (2000), 135–148. https://doi.org/10.1111/1468-0262.00096
Phelps, Edmond, “The Statistical Theory of Racism and Sexism,” *American Economic Review*, 62 (1972), 659–661.
Quah, John K.-H., and Bruno Strulovici, “Comparative Statics, Informativeness, and the Interval Dominance Order,” *Econometrica*, 77 (2009), 1949–1992. https://doi.org/10.3982/ECTA7583
Rashdall, Hastings, *The Universities of Europe in the Middle Ages: Volume I* (Oxford: Clarendon Press, 1895).
Roy, Andrew D., “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 3 (1951), 135–146. https://doi.org/10.1093/oxfordjournals.oep.a041827
Samuelson, Paul A., *Foundations of Economic Analysis* (Cambridge, MA: Harvard University Press, 1947).
Savage, James D., *Funding Science in America: Congress, Universities, and the Politics of the Academic Pork Barrel* (Cambridge: Cambridge University Press, 1999). https://doi.org/10.1017/CBO9780511625558
Scotchmer, Suzanne, *Innovation and Incentives* (Cambridge, MA: MIT Press, 2004).
Seglen, Per O., “The Skewness of Science,” *Journal of the American Society for Information Science*, 43 (1992), 628–638. https://doi.org/10.1002/(SICI)1097-4571(199210)43:9%3C628::AID-ASI5%3E3.0.CO;2-0
Siegel, Ron, “All-Pay Contests,” *Econometrica*, 77 (2009), 71–92. https://doi.org/10.3982/ECTA7537
Stephan, Paula, *How Economics Shapes Science* (Cambridge, MA: Harvard University Press, 2012).
Taylor, Curtis R., and Huseyin Yildirim, “Subjective Performance and the Value of Blind Evaluation,” *Review of Economic Studies*, 78 (2011), 762–794. https://doi.org/10.1093/restud/rdq005
Thomson, William, *How to Divide When There Isn’t Enough: From Aristotle, the Talmud, and Maimonides to the Axiomatics of Resource Allocation* (Cambridge: Cambridge University Press, 2019).
Upchurch, Anna Rosser, *The Origins of the Arts Council Movement: Philanthropy and Policy* (London: Palgrave Macmillan, 2016). https://doi.org/10.1057/978-1-137-46163-6

van Zwet, William R., *Convex Transformations of Random Variables* (Amsterdam: Mathematisch Centrum, 1964).
von, Hippel Ted, and Courtney von Hippel, "To Apply or Not to Apply: A Survey Analysis of Grant Writing Costs and Benefits," *PLoS One*, 10 (2015), e0118494.
Weinstein, Milton, and Richard Zeckhauser, "Critical Ratios and Efficient Allocation," *Journal of Public Economics*, 2 (1973), 147–157. https://doi.org/10.1016/0047-2727(73)90002-9
Westfall, Richard S., "Science and Patronage: Galileo and the Telescope," *Isis*, 76 (1985), 11–30. https://doi.org/10.1086/353735
Weisbrod, Burton, *Economics of Public Health* (Philadelphia: University of Pennsylvania Press, 1961). https://doi.org/10.9783/9781512808643
Wildavsky, Aaron, *The Politics of the Budgetary Process* (Boston: Little, Brown, 1964).
Zeckhauser, Richard, "Some Thoughts on the Allocation of Resources to BioMedical Research," U.S. Department of Health, Education and Welfare Occasional Paper No. 4, (U.S. Department of Health, Education and Welfare: Office of Planning and Evaluation, 1967).
Zuckerman, Harriet, and Robert K. Merton, "Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System," *Minerva*, 9 (1971), 66–100. https://doi.org/10.1007/BF01553188
Zunz, Olivia, *Philanthropy in America: A History* (Princeton, NJ: Princeton University Press, 2012).