

An efficient system to fund science: from proposal review to peer-to-peer distributions

Johan Bollen^{1,3,4} · David Crandall^{1,4} · Damion Junk¹ · Ying Ding^{1,3} · Katy Börner^{1,2,3,4}

Received: 17 March 2016
© Akadémiai Kiadó, Budapest, Hungary 2016

Abstract This paper presents a novel model of science funding that exploits the wisdom of the scientific crowd. Each researcher receives an equal, unconditional part of all available science funding on a yearly basis, but is required to individually donate to other scientists a given fraction of all they receive. Science funding thus moves from one scientist to the next in such a way that scientists who receive many donations must also redistribute the most. As the funding circulates through the scientific community it is mathematically expected to converge on a funding distribution favored by the entire scientific community. This is achieved without any proposal submissions or reviews. The model furthermore funds scientists instead of projects, reducing much of the overhead and bias of the present grant peer review system. Model validation using large-scale citation data and funding records over the past 20 years show that the proposed model could yield funding distributions that are similar to those of the NSF and NIH, and the model could potentially be more fair and more equitable. We discuss possible extensions of this approach as well as science policy implications.

Keywords Bibliometrics · Funding · Peer review · PageRank · Collective intelligence

Electronic supplementary material The online version of this article (doi:[10.1007/s11192-016-2110-3](https://doi.org/10.1007/s11192-016-2110-3)) contains supplementary material, which is available to authorized users.

✉ Johan Bollen
jbollen@indiana.edu

¹ School of Informatics and Computing, Indiana University, Bloomington, IN, USA

² Department of Information and Library Science, Indiana University, Bloomington, IN, USA

³ Indiana University Network Institute, Indiana University, Bloomington, IN, USA

⁴ Center for Complex Network and Systems Research, Indiana University, Bloomington, IN, USA

Introduction

Public agencies such as the US National Science Foundation (NSF) and the National Institutes of Health (NIH) award tens of billions of dollars in science funding annually. How can this money be distributed as efficiently as possible to best promote scientific innovation and productivity?

During 2015 alone, the NSF conducted 231,000 proposal reviews to evaluate 49,600 proposals that directly funded 350,000 people (researchers, postdocs, trainees, teachers, and students) (National Science Foundation 2015). Although considered the scientific gold standard (Jefferson 2006), grant peer review requires significant overhead costs (Southwick 2012; Roy 1985). Herbert et al. (2013) estimate that Australian researchers alone spent five centuries worth of research time preparing proposals. Extrapolating for population size, the US grant review system may incur even greater overall overhead costs. This situation may be no less favorable at the individual or institutional levels where faculty spend significant amounts of valuable time preparing and submitting research proposals. According to at least one study, faculty report to spend 42 % of their time attending to pre- and post-award administrative demands (FSC 2007). Principal investigators (PI) and co-PIs may spend 116 and 55 h per proposal respectively. As a result, the monetary value of any subsequent grant may be greatly diminished when taking into account all preparation, submission, reviewing, as well as professional and personal costs (Herbert et al. 2014), even to the point that applying for grants might eventually become, on balance, a poor investment for universities and faculty alike.

Peer review may furthermore be subject to biases, inconsistencies, and oversights that are difficult to remediate (Gura 2002; Suresh 2012; Wessely 1998; Cicchetti 1991; Cole and Cole 1981; Azoulay et al. 2012; Horrobin 1996; Editorial 2013; Ioannidis and Nicholson 2012; Smith 2006; Marsh et al. 2008; Bornmann et al. 2007; Johnson et al. 2008). These issues have led some to propose less costly, and possibly more reliable and effective alternatives, such as the random distribution of funding (Geard and Noble 2010) or person-directed funding that does not involve proposal preparation and review (Azoulay et al. 2009). Proposals to reform funding systems have ranged from incremental changes to the peer review process including careful selection of reviewers (Marsh et al. 2008) and post-hoc normalization of reviews (Johnson et al. 2008), to more radical proposals such as opening the proposal review process to the entire online population (Editor 2010) or removing human reviewers altogether by allocating funds equally, randomly, or through an objective performance measure (Roy 1985).

Here we investigate a new class of funding models in which all scientists individually participate in the allocation of research funding. All participants receive an equal portion of all yearly funding, but they are then required to anonymously donate a fraction of everything that they have received to peers. Funding thus flows from one participant to the next, each acting as if he or she were a funding agency themselves. These distributed systems incorporate the opinions of the entire scientific community, but in a highly-structured framework that encourages fairness, robustness, and efficiency.

We use large-scale citation data (37 million articles, 770 million citations) as a proxy for how researchers might distribute their funds in a large-scale simulation of the proposed system. Model validation suggests that such a distributed system for science yields funding patterns similar to existing NIH and NSF distributions, but it may do so at much lower overhead while exhibiting a range of other desirable features. Our results indicate that self-correcting mechanisms in scientific peer evaluation can yield an efficient and fair

distribution of funding. The proposed model can be applied in many other situations in which top-down or bottom-up allocation of public resources is either impractical or undesirable, e.g., public investments, distribution chains, and shared resource management.

Methods

In the proposed system, all scientists are given an equal and unconditional base amount of yearly funding. Each year, however, they are also required to distribute a given percentage of their funding to other scientists whom they feel would make best use of the money (Fig. 1). Each scientist thus receives funds from two sources: the fixed based amount they receive unconditionally and the amounts they receive from other scientists. As scientists donate a fraction of their total funding to other scientists each year, funding moves from one scientist to the next. Everyone is guaranteed the base amount, but larger amounts will accumulate with scientists whom most scientists believe will make best use of the funding.

For example, suppose the base funding amount is set to \$100,000. This roughly corresponds to the entire NSF budget in 2010 divided by the total number of senior researchers it funded (National Science Foundation 2009). If the required donation fraction is set to $F = 0.5$, i.e., 50 %, scientist K receives her yearly base amount of \$100,000. In

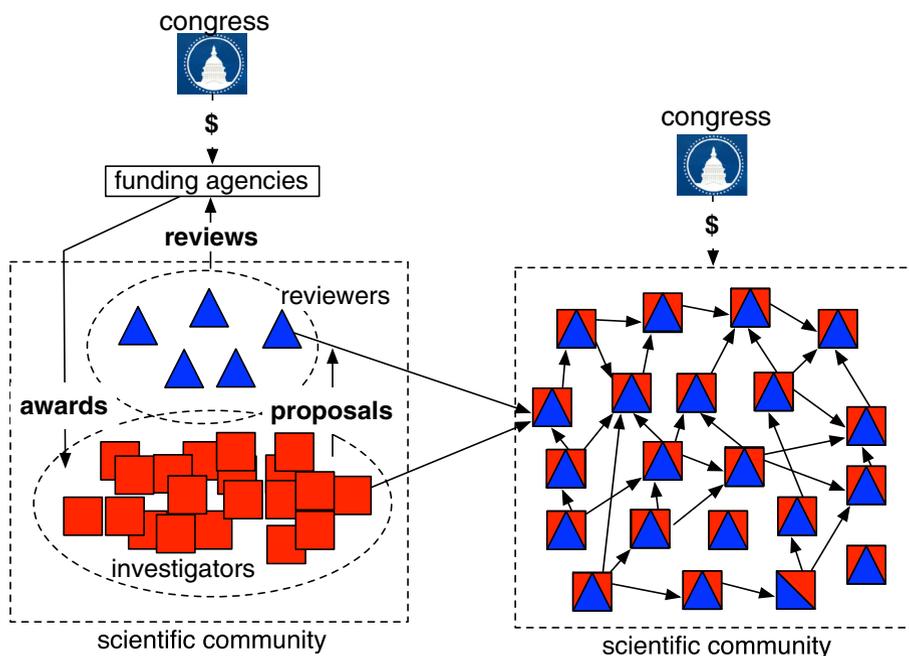


Fig. 1 Illustrations of existing (*left*) and the proposed (*right*) funding systems, with reviewers marked with triangles and investigators marked by squares. In most current funding models like those used by NSF and NIH, investigators write proposals in response to solicitations from funding agencies, these proposals are reviewed by small panels, and funding agencies use these reviews to help make funding decisions, providing awards to some investigators. In the proposed system, all scientists are both investigators and reviewers: every scientist receives a fixed amount of funding from the government and other scientists but is required to redistribute some fraction of it to other investigators

addition she may receive \$200,000 from other scientists in 2012, bringing her total funding to \$300,000. In 2013, K can spend half of that total, i.e., \$150,000, on her own research program, but must donate the other half to other scientists for their use in 2014. Rather than painstakingly submitting project proposals, K and her colleagues only need to take a few minutes of their time each year to log into a centralized website and enter the names of the scientists they choose to donate to and how much each should receive.

More formally, suppose that a funding agency maintains an account for each of the N qualified scientists (chosen according to some criteria such as academic appointment status or recent research productivity), to whom we assign unique identifiers in $[1, N]$. Let $O_{i \rightarrow j}^t$ denote the amount of money that scientist i gave to scientist j in year t . The amount of funding A each scientist receives in year t is equal to the base funding B from the government plus the contributions from other scientists,

$$A_i^t = B + \sum_{j \in [1, N]} O_{j \rightarrow i}^{t-1}.$$

We require that every scientist gives away a fraction F of their funds each year,

$$\sum_{j \in [1, N]} O_{i \rightarrow j}^t = (F)A_i^t \quad \forall i \in N,$$

while they are allowed to spend the remaining money on their research activities. Taken together, these two equations provide a “recursive” definition of the behavior of the overall funding system. A similar formulation has been used to rank webpages by transferring influence from one page to the next (Brin and Page 1998), as well as to rank scientific journals (Bollen et al. 2006), and author “prestige” (Ding et al. 2009).

This simple, highly distributed process yields surprisingly sophisticated behavior at a global level. First, respected and highly-productive scientists are likely to receive a comparatively large number of donations. They must in turn distribute a fraction of this total to others; their high status among scientists thus affords them both greater funding and greater influence over how funding is distributed. Second, as the priorities and preferences of the scientific community change over time, the flow of funding will change accordingly. Rather than converging on a stationary distribution, the system will dynamically adjust funding levels as scientists collectively assess and re-assess each others’ merits.

Results

How would funding decisions made by this system compare to the gold standard of peer review? No funding system will be optimal since research outcomes and impact are difficult to predict in advance (Myhrvold 1998). At the very least, one might hope that the outcome of the proposed system would approximate those of existing funding systems.

To investigate whether this minimal criterion might be satisfied, we conducted a large-scale, agent-based simulation to test how the proposed funding system might operate in practice. For this simulation, we used citations as a proxy for how each scientist might distribute funds in the proposed system. We extracted 37 million academic papers and their 770 million references from Thomson-Reuters’ 1992 to 2010 Web of Science (WoS) database. We converted this data into a citation graph by matching each reference with a paper in the database by comparing year, volume, page number, and journal name, while

allowing for some variation in journal names due to abbreviations. About 70 % of all references could be matched back to a manuscript within the WoS data itself.

From the matching 37 million papers, we retrieved 4,195,734 unique author names; we took the 867,872 names who had authored at least one paper per year in any consecutive 5 years of the period 2000–2010 to be in the set of qualified scientists for the purpose of our study. For each pair of authors, we determined the number of times one had cited the other in each year of our citation data (1992–2010). We also retrieved NIH and NSF funding records for the same period, a data set which provided 347,364 grant amounts for 109,919 unique scientists (Rowe et al. 2007).

We then ran our simulation beginning in the year 2000, in which everyone was given a fixed budget of $B = \$100,000$. We simulated the system by assuming that all scientists would distribute their funding in proportion to their citations over the prior 5 years. For example, if a scientist cited paper A three times and paper B two times over the last 5 years, then she would distribute three-fifths of her budget equally among the authors of A, and two-fifths amongst the authors of B.

Importantly, we stress that we are merely using citation data as a proxy for how scientists might distribute their funding for purposes of simulation, and are not proposing that actual funding decisions be made on the basis of citation analysis. Of course, scientists cite papers for a variety of reasons, not all of which indicate positive endorsement or influence on their work (MacRoberts and MacRoberts 2010). Nevertheless, although this proxy is an imperfect prediction of scientists' local funding distribution decisions, it permits a large-scale simulation that provides an initial indication of how the overall system may operate in practice.

The results of our simulation suggest that the resulting funding distribution would be heavy-tailed (Fig. 2a), and similar in shape to the actual funding distribution of NSF and NIH for the period 2000–2010 if $F \simeq 0.5$. As expected, the redistribution fraction F controls the shape of the distribution, with low values creating a nearly uniform distribution (less redistribution) and high values creating a highly biased distribution (more redistribution). This suggests that the value of F could be changed on the basis of policy objectives to control the level of funding inequality.

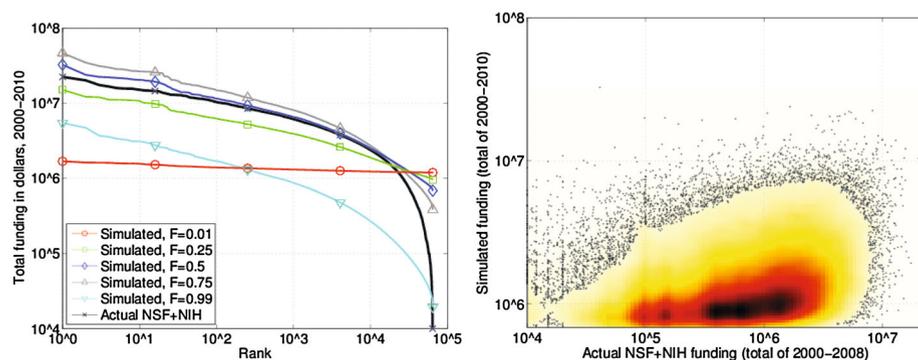


Fig. 2 Results of the distributed funding system simulation for 2000–2010. **a** The general shape of the funding distribution is similar to that of actual historical NSF and NIH funding distribution. The shape of the distribution can be controlled by adjusting F , the fraction of funds that scientists must give away each year. **b** On a per-scientist basis, simulated funding from our system (with $F = 0.5$) is correlated with actual NSF and NIH funding (Pearson $R = 0.2683$ and Spearman $\rho = 0.2999$)

Finally, we used a very conservative (and simple) heuristic to match the author names from our simulation results to those listed in the actual NSF and NIH funding records: we simply normalized all names to consist of a last name, a first name, and middle initials, and then required exact matches between the two sets. This conservative heuristic yielded 65,610 matching scientist names. For each scientist we compared their actual NSF and NIH funding for 2000–2010 to the amount of funding predicted by simulation of the proposed system, and we found that the two were correlated with Pearson $R = 0.268$ and Spearman $\rho = 0.300$ (Fig. 2b).

Discussion and conclusions

These results suggest that our proposed system would lead to funding distributions that are highly similar in shape and individual level of funding to those of the NIH and NSF, if scientists are compelled to redistribute 50 % of their funding each year—but at a fraction of the time and cost of the current peer-review system.

We note that the ability to mimic or reproduce the shape of the existing funding distribution is not a *conditio sine qua non*. It is certainly possible, and perhaps even likely given the underlying mathematics, that the proposed system yields funding distributions that are dissimilar from those obtained by the grant peer review system, and could potentially be more valid, equitable, and supportive of scientific innovation. This, however, cannot be determined from our simulation, which used citation behavior as a proxy for actual funding decisions.

As shown in Fig. 2, our results indicate very high levels of funding inequality in the present distribution of NIH and NSF funding, indicating that few individuals (or projects) receive very high amounts of funding while most receive rather low amounts of funding. Determining whether this situation promotes scientific innovation or not falls outside the scope of this work, but the proposed system provides a straightforward mechanism to create higher or lower levels of funding equality. At high levels of redistribution, the system becomes more strongly focused on merit as scientists retain less of their base amount and become more strongly dependent on donations from other scientists. This may lead to increasing levels of funding inequality. At low levels of redistribution, scientists retain more of the base amount and receive fewer donations. In other words, the system becomes less meritorious and more egalitarian, leading to higher levels of funding equality. The redistribution factor (F) thus provides policy makers with powerful new leverage to render the distribution of science funding more equitable (less redistribution) or more meritorious (more redistribution) than is presently the case.

The proposed framework would fund *people* (Gilbert 2009) instead of funding *projects*. Changes to the present grant peer review system have not significantly reduced the average age of first time grant recipients (Kaiser 2008), but the proposed system supports all scientists equally regardless of age or career development stage. Since every scientist receives an unconditional base amount every year, early career scientists could focus on building their research programs rather than spending valuable research time and resources to acquire funding (Editor 2011).

In general, our system would significantly reduce the amount of time scientists spend preparing and submitting proposals, freeing more time for scientific discovery and innovation. It also introduces incentives for a more open communication of scientific results and research plans. To receive donations, scientists would have to actively communicate

the value of their work to the larger scientific community and to the public. A strong commitment to clear communication, open science, and transparency might prove to be more effective than pursuing traditional tokens of academic merit, e.g., publications and citations. Conferences, workshops, journals, and publishers may have to change their *modus operandi* accordingly.

Of course, funding agencies and governments may still wish to play a directive role, e.g., to encourage advances in certain areas of national interest or to foster diversity. This could be included in the outlined system in a number of straightforward ways. Most simply, traditional peer-reviewed, project-based funding could be continued in parallel, using the proposed system to distribute only a portion of public science funding. Traditional avenues may also be needed to fund projects that develop or rely on large scientific tools and infrastructure. Alternatively, funding agencies could vary the base funding rate B across different constituencies, allowing them to temporarily inject more money into certain disciplines or certain underrepresented groups. The system could also include explicit temporal dampening to prevent large funding changes from year to year, e.g., by allowing scientists to save or accumulate portions of their funding.

In practice, the system will require Conflict of Interest rules similar to the ones that presently keep grant peer-review fair and unbiased. For example, scientists might be prevented from funding advisors, advisees, close collaborators, co-authors, or researchers at their own institution. The interface of an online funding platform might automatically preclude such donations. Donations must furthermore be kept confidential in order to prevent groups of people from colluding to affect the global funding distribution. Such collusion might however be easily detectable since large-scale, objective, but anonymous donation data will be available to analysts and policy-makers. In fact, being able to detect “gaming” may be another major advantage over the current system.

In summary, peer-review of funding proposals has served science well for decades, but funding agencies may want to consider alternative approaches to public science funding that build on theoretical advances and leverage modern technology and big data analytics. The system proposed here requires a fraction of the costs associated with peer review. The potential savings of financial as well as human resources could be used to better identify targets of opportunity, to translate scientific results into products and jobs, and to help communicate scientific and technological advances to the public and to policy makers.

Acknowledgments The authors acknowledge the generous support of the National Science Foundation under Grant SBE #0914939 and SMA #1636636, the National Institutes of Health under Grants #P01AG039347 and #U01GM098959, and the Andrew W. Mellon Foundation. We also thank the Los Alamos National Laboratory Research Library, the LANL Digital Library Prototyping and Research Team, Thomson-Reuters, and the Cyberinfrastructure for Network Science Center at Indiana University for furnishing the data employed in this analysis. The authors thank Marten Scheffer (Wageningen University) for his extensive feedback on our work and his support of in vivo implementations.

Compliance with ethical standards

Competing interest The authors declare that they have no competing financial interest.

References

Azoulay, P., Zivin, J. S. G., & Manso, G. (2009). Incentives and creativity: Evidence from the academic life sciences. Working Paper 15466, National Bureau of Economic Research. doi:10.3386/w15466. <http://www.nber.org/papers/w15466>.

- Azoulay, P., Zivin, J. S. G., & Manso, G. (2012). *NIH peer review: Challenges and avenues for reform*. National Bureau of Economic Research Working Paper 18116.
- Bollen, J., Rodriguez, M. A., & Van de Sompel, H. (2006). Journal status. *Scientometrics*, 69(3), 669.
- Bormmann, L., Mutz, R., & Daniel, H. D. (2007). Gender differences in grant peer review: A meta-analysis. *Journal of Informetrics*, 1, 226.
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30, 107.
- Cicchetti, D. V. (1991). The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences*, 14(1), 119–135.
- Cole, G. S. S., & Cole, J. R. (1981). Chance and consensus in peer review. *Science*, 214, 881.
- Ding, Y., Yan, E., Frazho, A., & Caverlee, J. (2009). PageRank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology*, 60(11), 2229.
- Editor (2010). Calm in a crisis. *Nature*, 468, 1002.
- Editor (April 25 2011). *Dr. No money: The broken science funding system*. Scientific American.
- Editorial (2013). Twice the price. *Nature*, 493(7434).
- F.S.C. of the Federal Demonstration Partnership (2007). *A profile of federal grant administrative burden among federal demonstration partnership faculty*. Tech. rep. http://www.iscintelligence.com/archivos_subidos/usfacultyburden_5.pdf.
- Geard, N., & Noble, J. (2010) Modelling academic research funding as a resource allocation problem. In *3rd world congress on social simulation*.
- Gilbert, N. (2009). Wellcome trust makes it personal in funding revamp: People not projects are the focus of longer-term grants. *Nature (Online)*, 462(145), 145. doi:10.1038/462145a.
- Gura, T. (2002). Scientific publishing: Peer review, unmasked. *Nature*, 416(6878), 258.
- Herbert, D. L., Barnett, A. G., & Graves, N. (2013). Funding: Australia's grant system wastes time. *Nature*. doi:10.1038/495314d.
- Herbert, D. L., Coveney, J., Clarke, P., Graves, N., & Barnett, A. G. (2014). The impact of funding deadlines on personal workloads, stress and family relationships: A qualitative study of Australian researchers. *BMJ Open*. doi:10.1136/bmjopen-2013-004462.
- Horrobin, D. (1996). Peer review of grant applications: A harbinger for mediocrity in clinical research? *The Lancet*, 348, 1293.
- Ioannidis, J. M., & Nicholson, J. P. A. (2012). Research grants: Conform and be funded. *Nature*, 492, 34. doi:10.1038/492034a.
- Jefferson, T. (2006). Quality and value: Models of quality control for scientific research. *Nature*. doi:10.1038/nature05031.
- Johnson, V. E. (2008). Statistical analysis of the National Institutes of Health peer review system. *Proceedings of the National Academy of Sciences*, 105, 11076.
- Kaiser, J. (2008). Zerhouni's parting message: Make room for young scientists. *Science*, 322(5903), 834.
- MacRoberts, M. H., & MacRoberts, B. R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science*, 61(1), 1–12.
- Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist*, 63, 160.
- Myhrvold, N. (1998). Supporting science. *Science*, 282, 621.
- National Science Foundation (2009). *Fiscal year 2010 budget request to congress*.
- National Science Foundation (2015). *Fy2015 agency financial report*. NSF report nsf16002. http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf16002.
- Rowe, G. L., Burgoon, S., Burgoon, J., Ke, W., & Börner, K. (2007). The scholarly database and its utility for scientometrics research. In *Proceedings of the 11th international conference on scientometrics and informetrics* (pp. 457–462).
- Roy, R. (1985). Funding science: The real defects of peer review and an alternative to it. *Science, Technology, and Human Values*, 10, 73.
- Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, 99, 178.
- Southwick, F. (2012). Academia suppresses creativity. *The Scientist*.
- Suresh, S. (2012). Research funding: Global challenges need global solutions. *Nature*, 490, 337.
- Wessely, S. (1998). Peer review of grant applications: What do we know? *The Lancet*, 352, 301.