Modeling effects of low funding rates on innovative research

Pawel Sobkowicz

March 8, 2016

Introduction

Peer review is the cornerstone of modern science: from the publication process to the evaluation of funding applications. There are, however fundamental differences between the role of the peer review in the review of publications and in the evaluation of funding requests:

Introduction

Peer review is the cornerstone of modern science: from the publication process to the evaluation of funding applications. There are, however fundamental differences between the role of the peer review in the review of publications and in the evaluation of funding requests:

- In publishing, the reviewers evaluate concrete results, in grant applications they evaluate promises;

Introduction

Peer review is the cornerstone of modern science: from the publication process to the evaluation of funding applications. There are, however fundamental differences between the role of the peer review in the review of publications and in the evaluation of funding requests:

- In publishing, the reviewers evaluate concrete results, in grant applications they evaluate promises;
- Negative decision of a publication submission is almost never a catastrophe: there are so many journals around. On the other hand, reviewer's decisions leading to a lack of funding may kill someone's career (and frequently do).

Introduction

Peer review is the cornerstone of modern science: from the publication process to the evaluation of funding applications. There are, however fundamental differences between the role of the peer review in the review of publications and in the evaluation of funding requests:

- In publishing, the reviewers evaluate concrete results, in grant applications they evaluate promises;
- Negative decision of a publication submission is almost never a catastrophe: there are so many journals around. On the other hand, reviewer's decisions leading to a lack of funding may kill someone's career (and frequently do).

Our goal: an agent based model that uncovers the negative effects of the current reliance on the competitive grant schemes in science funding.

Some quotes

At first glance the notion of "excellence through competition" seems reasonable. The idea is relatively easy to sell to politicians and the general public. [...] On the practical side, **the net result of the heavy-duty "expert-based" peer review system is that more often than not truly innovative research is suppressed.**
Furthermore, the secretive nature of **the funding system efficiently turns it into a self-serving network operating on the principle of an "old boys' club."**
A Berezin, The perils of centralized research funding systems, 1998

Some quotes

Diversity – which is essential, since experts cannot know the source of the next major discovery – is not encouraged. [. . .] **The projects funded will not be risky, brilliant, and highly innovative since such applications would inevitably arouse broad opposition from the administrators, the reviewers, or some committee members.** *[. . .] In the UK (and probably elsewhere), we are not funding worthless research. But we are funding research that is fundamentally pedestrian, fashionable, uniform, and second-league.*

D F Horrobin, Peer review of grant applications: a harbinger for mediocrity in clinical research?, 1996

Some quotes

Further cohort studies of unfunded proposals are needed. Such studies will, however, always be difficult to interpret – **do they show how peer review prevents resources from being wasted on bad science, or do they reveal the blinkered conservative preferences of senior reviewers who stifle innovation and destroy the morale of promising younger scientists?** *We cannot say.*

S Wessely, Peer review of grant applications: what do we know?, 1998

Model assumptions

- We start with N_P proposals are submitted each year, with starting $N_P = 2000$ and 2% growth each year.

Model assumptions

- We start with N_P proposals are submitted each year, with starting $N_P = 2000$ and 2% growth each year.
- We assume a lognormal distribution of innovation value $V(P)$ of proposals P.

Model assumptions

- We start with N_P proposals are submitted each year, with starting $N_P = 2000$ and 2% growth each year.
- We assume a lognormal distribution of innovation value $V(P)$ of proposals P.
- Only a small fraction (say, 20%) of the proposals get funded.

Model assumptions

- We start with N_P proposals are submitted each year, with starting $N_P = 2000$ and 2% growth each year.
- We assume a lognormal distribution of innovation value $V(P)$ of proposals P.
- Only a small fraction (say, 20%) of the proposals get funded.
- Out of the rejected ones, 60% are resubmitted with the same innovativeness value, 40% drop out, and are replaced by new proposals/researchers.

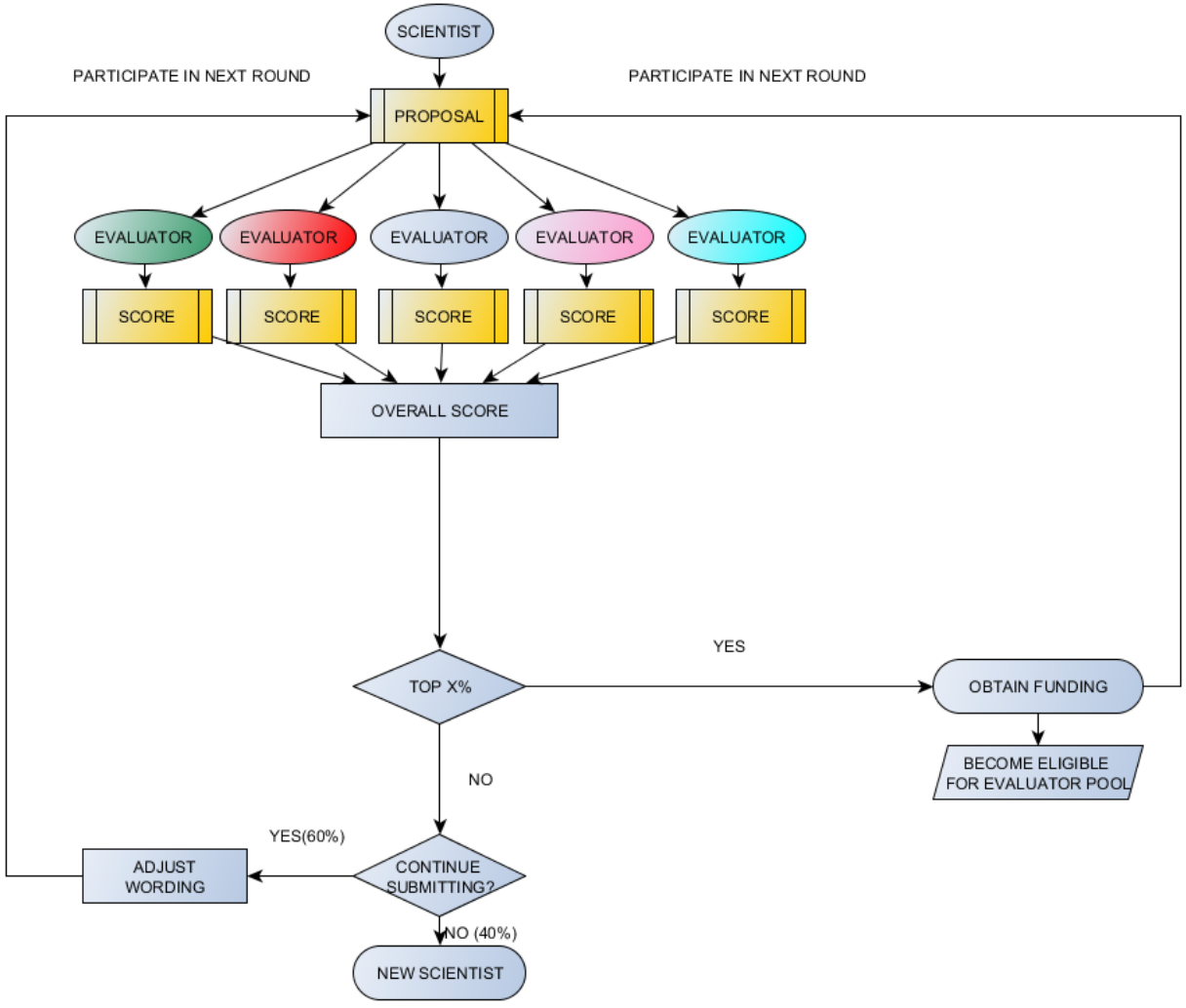
Model assumptions

- We start with N_P proposals are submitted each year, with starting $N_P = 2000$ and 2% growth each year.
- We assume a lognormal distribution of innovation value $V(P)$ of proposals P.
- Only a small fraction (say, 20%) of the proposals get funded.
- Out of the rejected ones, 60% are resubmitted with the same innovativeness value, 40% drop out, and are replaced by new proposals/researchers.
- Selection is done by groups of N_E (5) evaluators, drawn randomly from a pool of experts $\mathcal{R}$ of size N_X (300).

Model assumptions

- We start with N_P proposals are submitted each year, with starting $N_P = 2000$ and 2% growth each year.
- We assume a lognormal distribution of innovation value $V(P)$ of proposals P.
- Only a small fraction (say, 20%) of the proposals get funded.
- Out of the rejected ones, 60% are resubmitted with the same innovativeness value, 40% drop out, and are replaced by new proposals/researchers.
- Selection is done by groups of N_E (5) evaluators, drawn randomly from a pool of experts $\mathcal{R}$ of size N_X (300).
- In the **ideal world** case every evaluator would assign the proposal a score equal to its innovation value $S(P, E) = V(P)$ and only the proposals with topmost scores get funded.

Process flow – ideal case



Non-ideal world

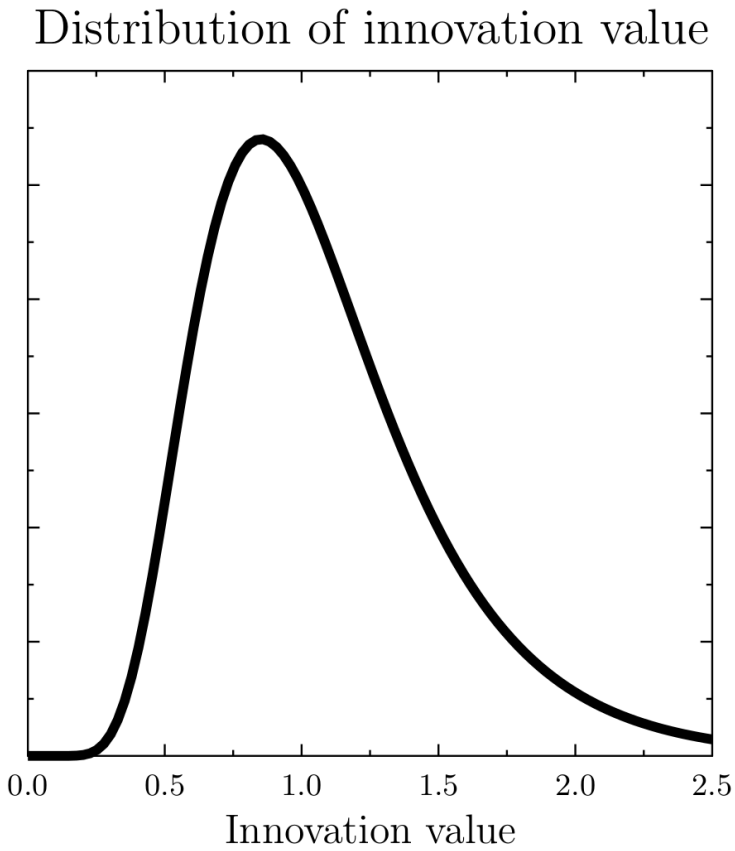
- Every evaluator suffers from limitations of his/her own innovativeness. Evaluator's own innovativeness acts thus as a tolerance filter for the evaluated proposals.
- Moreover, there is inevitable 'noise' in the system, which further decreases the accuracy of scoring.
- Lastly, many competitions, in addition to evaluation of proposals, include additional scores for the researcher/team quality, usually measured by their *past successes* . . .

Non-ideal world

- Every evaluator suffers from limitations of his/her own innovativeness. Evaluator's own innovativeness acts thus as a tolerance filter for the evaluated proposals.
- Moreover, there is inevitable 'noise' in the system, which further decreases the accuracy of scoring.
- Lastly, many competitions, in addition to evaluation of proposals, include additional scores for the researcher/team quality, usually measured by their *past successes* . . . in getting grants. Leading directly to the Matthew effect.

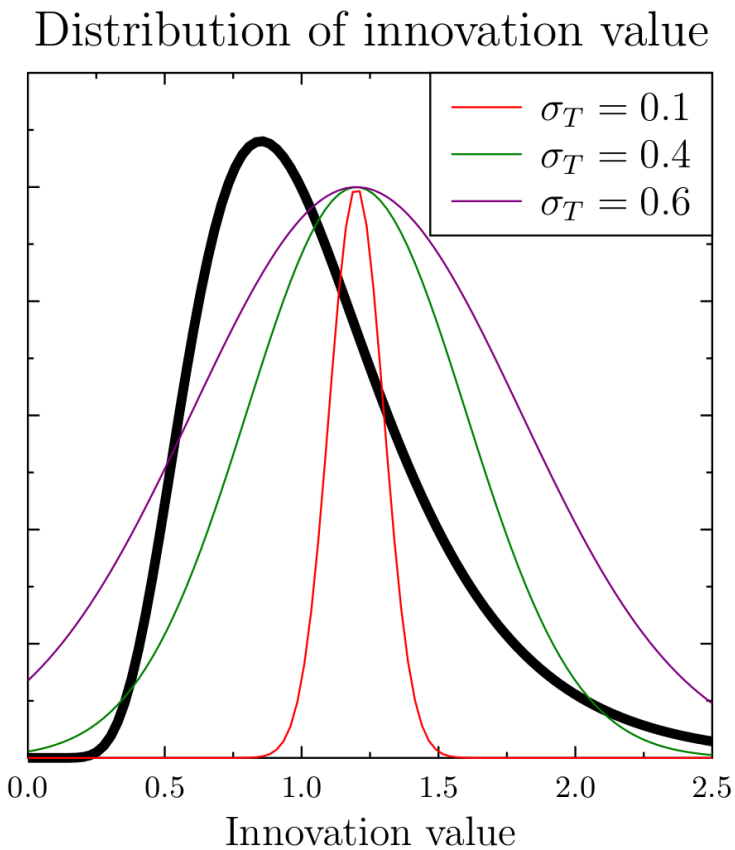
Tolerance filter in action

We start with the ‘raw’ lognormal distribution of the innovation values of the proposals



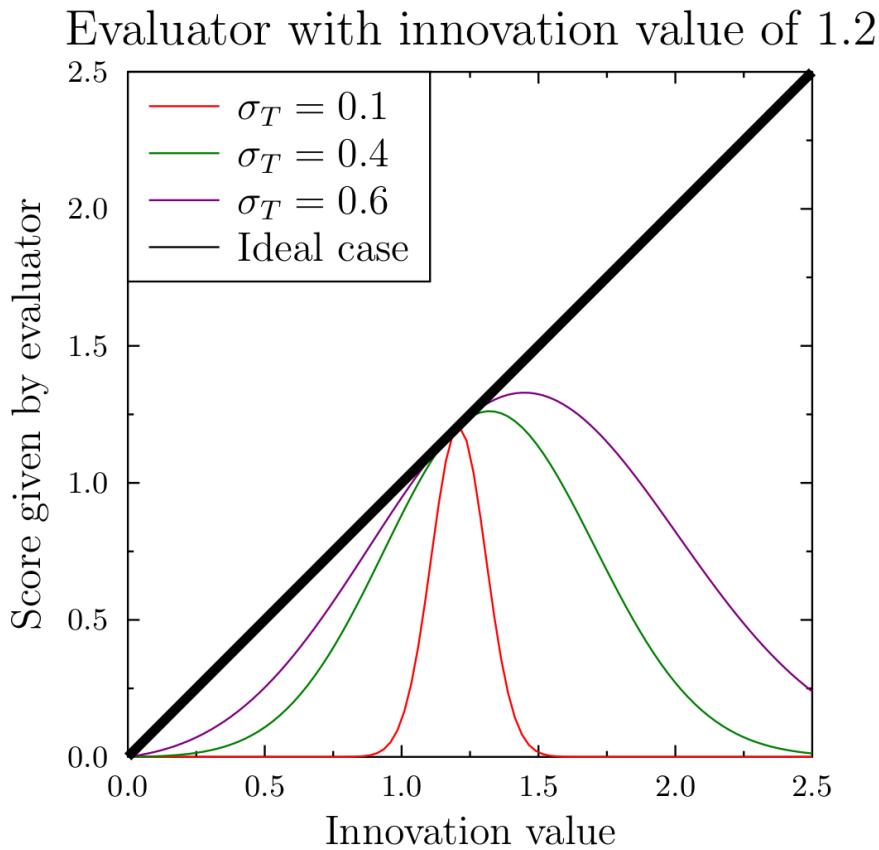
Tolerance filter in action

The filter example: the evaluator has innovativeness of 1.2 and three values of the tolerance σ_T.



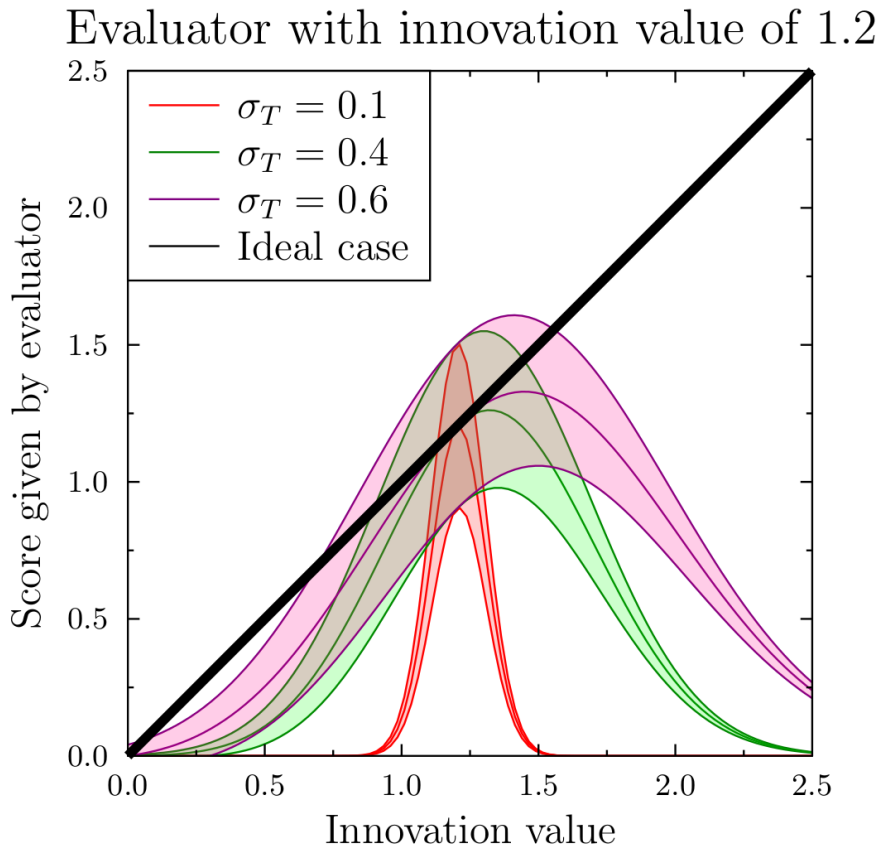
Tolerance filter in action

The resulting scores given by the evaluator. Horizontal axis: true innovation value, vertical axis: score.

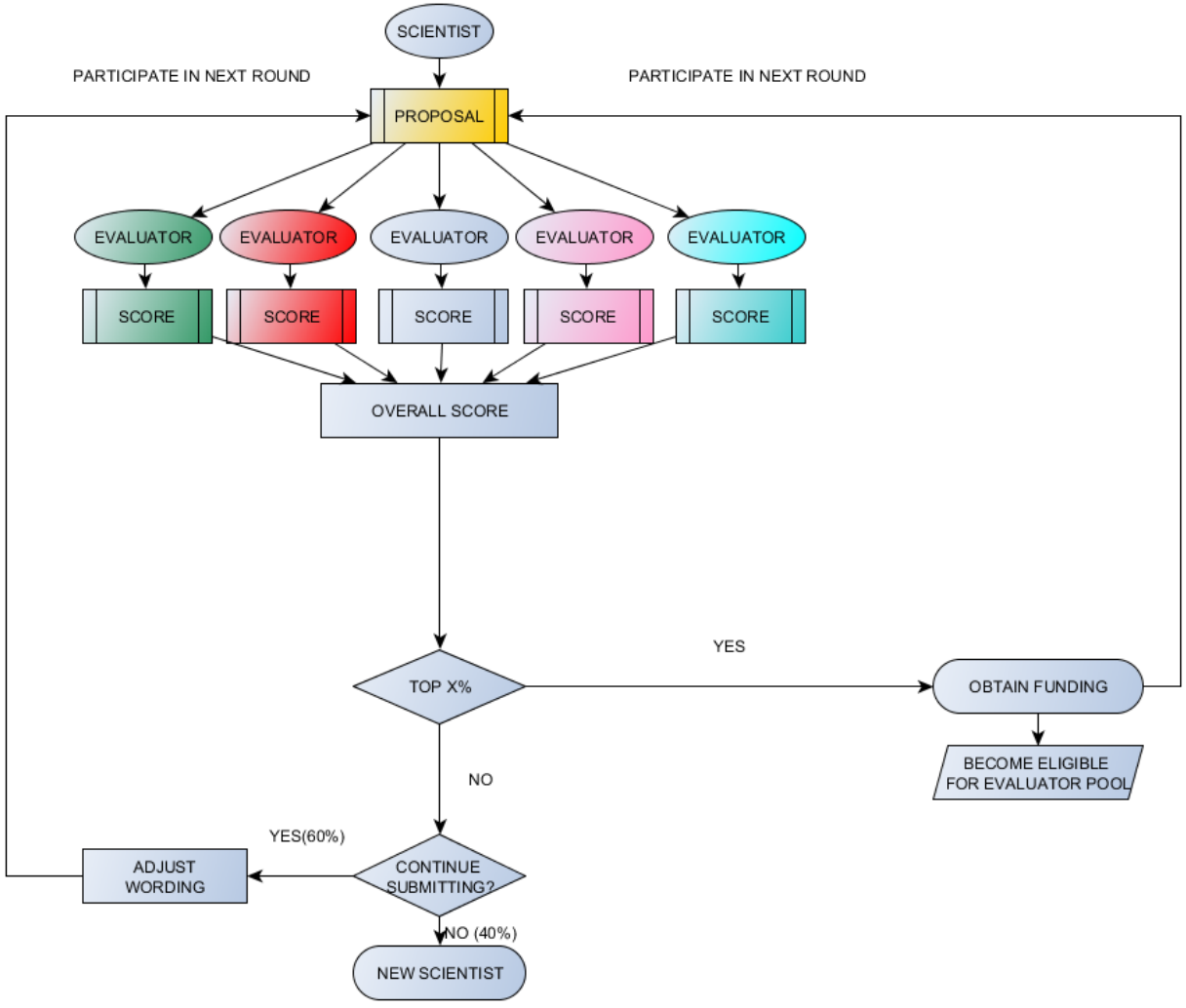


Tolerance filter in action

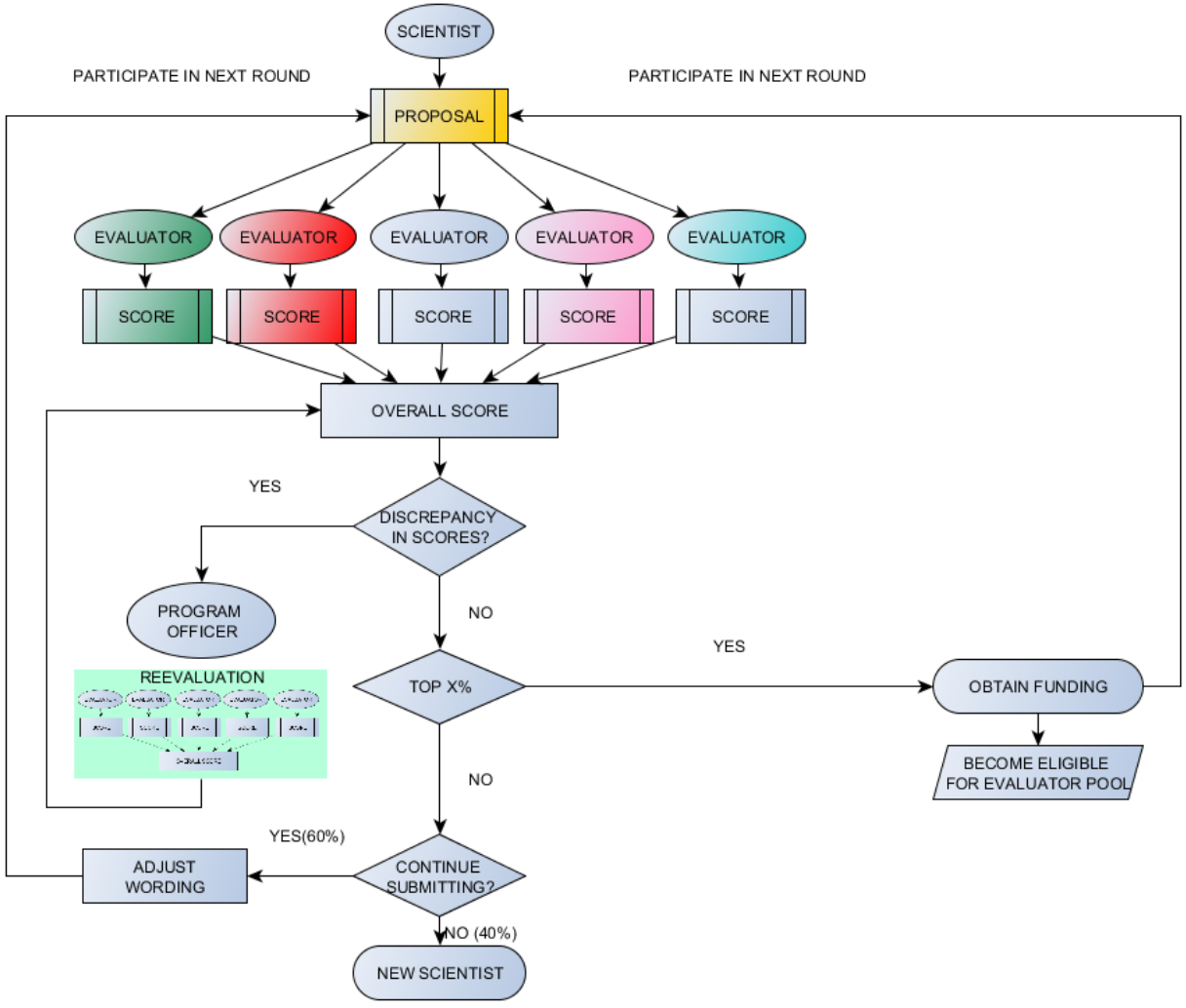
The resulting scores given by the evaluator. This time some 'noise' has been added to the evaluation process.



Process flow – non-ideal case



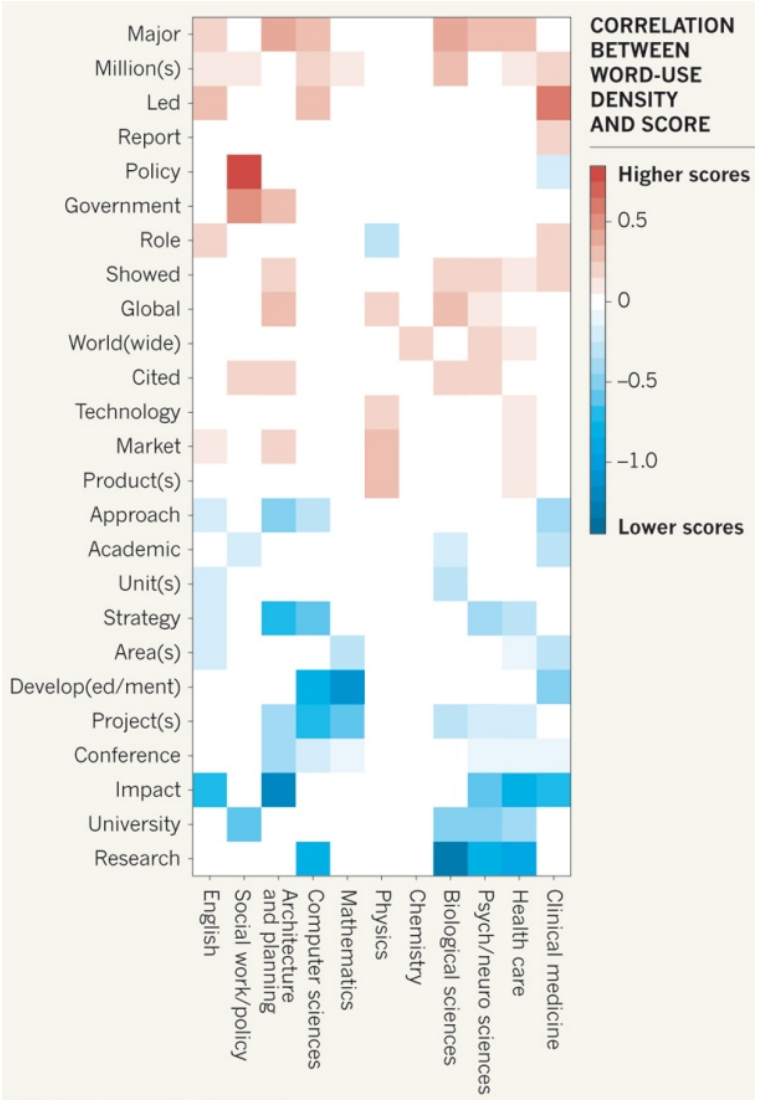
Process flow – with re-evaluation



Process flow – adjustment of proposals

The use of currently fashionable buzzwords will make proposals more alike: converging on the mean value, regardless of the actual innovation. And yes, there are magic words, and anyone can use them...

Van Noorden, R., *Seven thousand stories capture impact of science*. Nature, 2015, 518(7538), p.150.

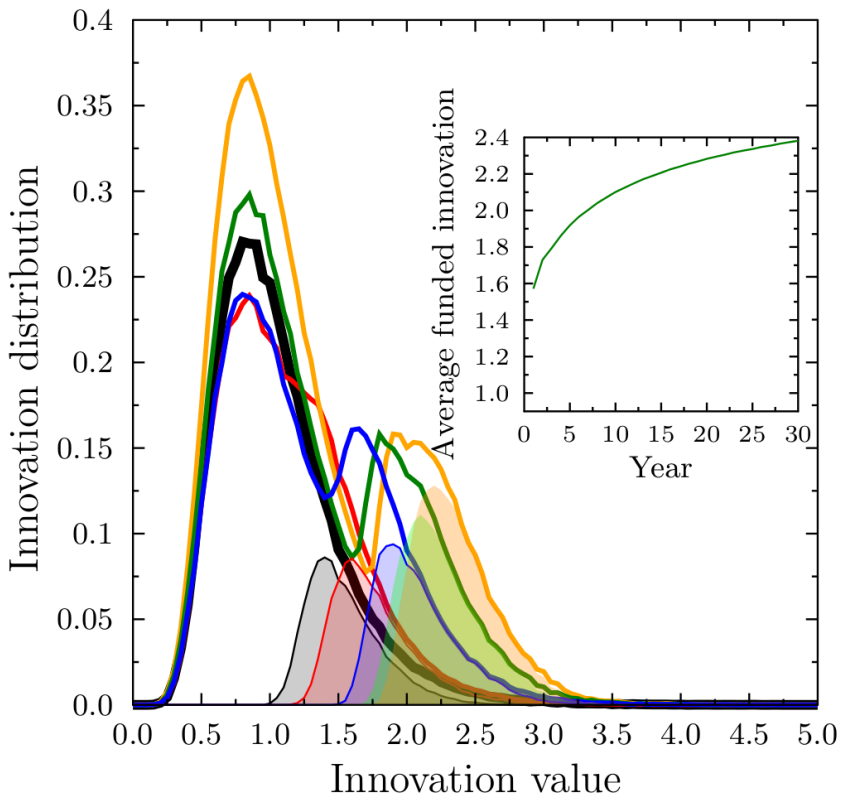


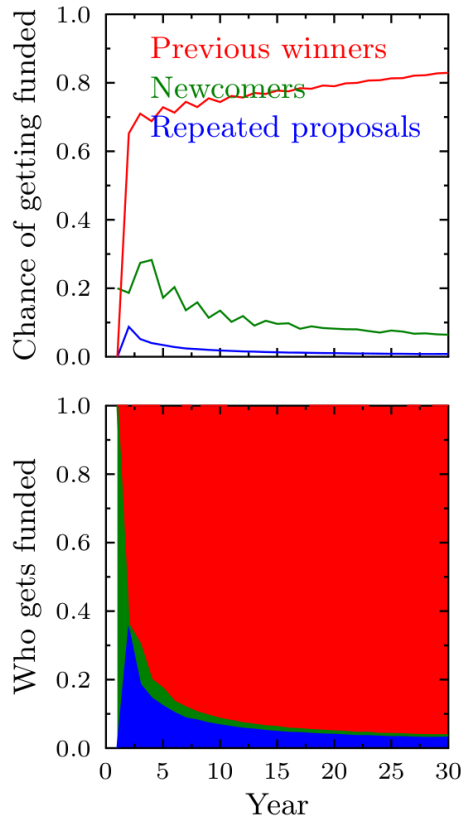
Model results in various circumstances

Ideal case. No re-evaluation.

High tolerance $\sigma_T = 1.0$. Noise ± 0.3.

Repeated submissions use more of the current 'newspeak'.



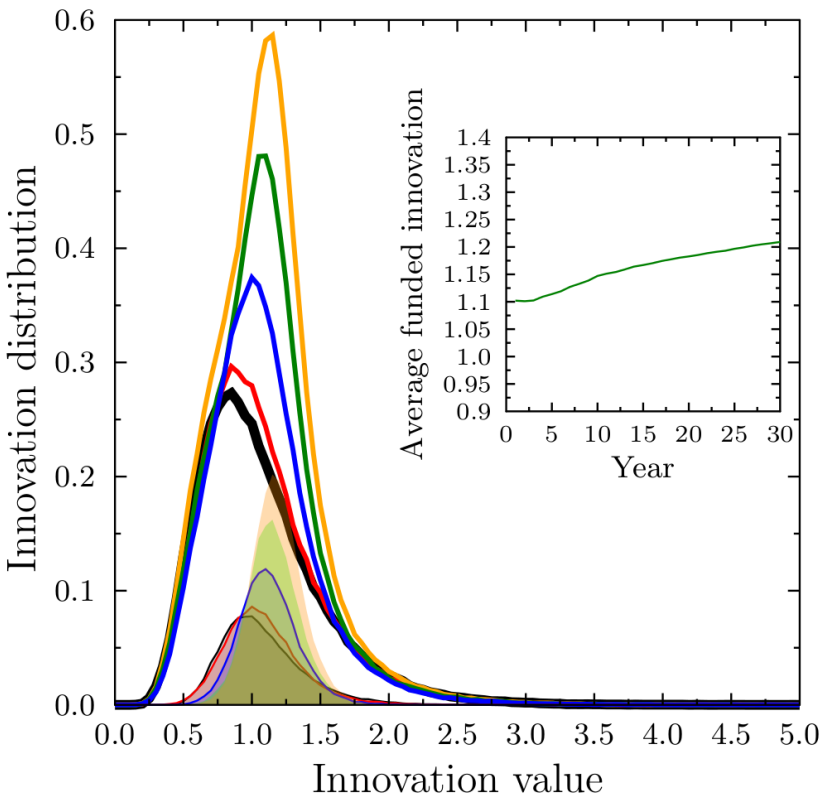


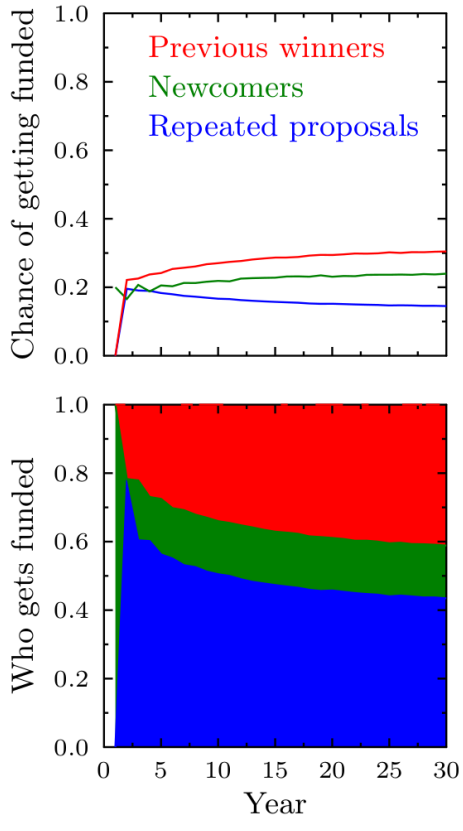
Model results in various circumstances

No previous success bonus. No re-evaluation.

Low tolerance $\sigma_T = 0.1$. Noise ± 0.3.

Repeated submissions use more of the current 'newspeak'.



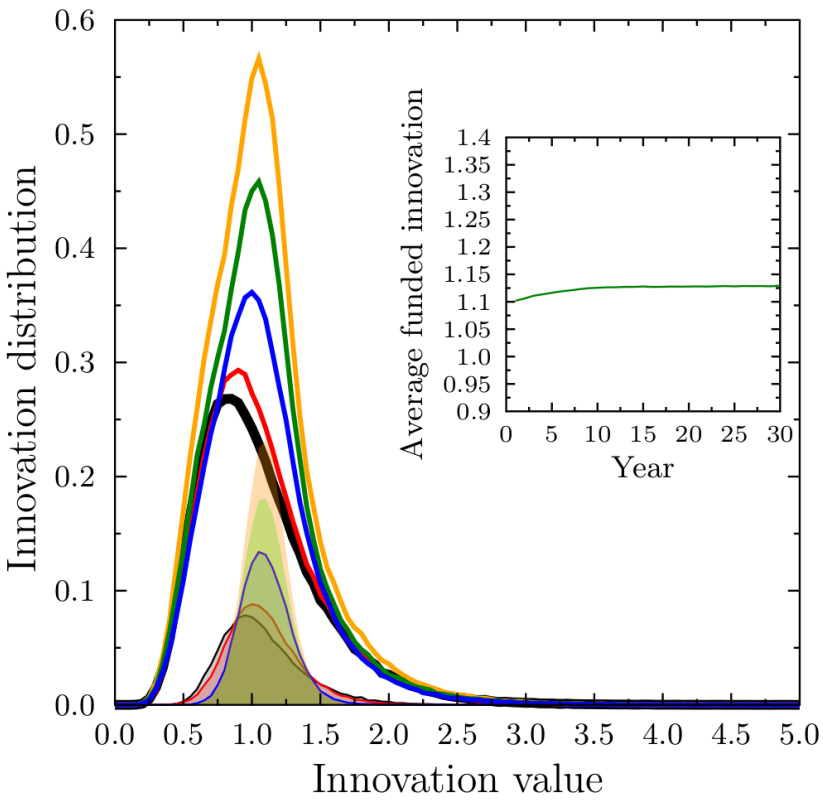


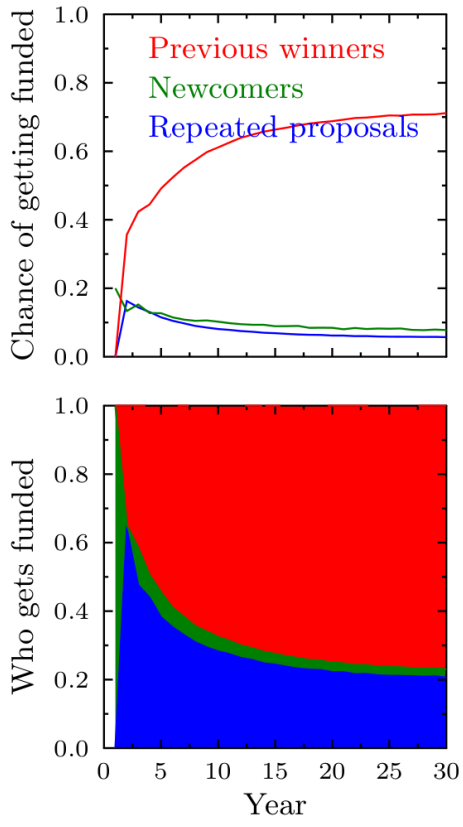
Model results in various circumstances

Bonus for previous succeses (0.1 per evaluation). No re-evaluation.

Low tolerance $\sigma_T = 0.1$. Noise ± 0.3.

Repeated submissions use more of the current 'newspeak'.



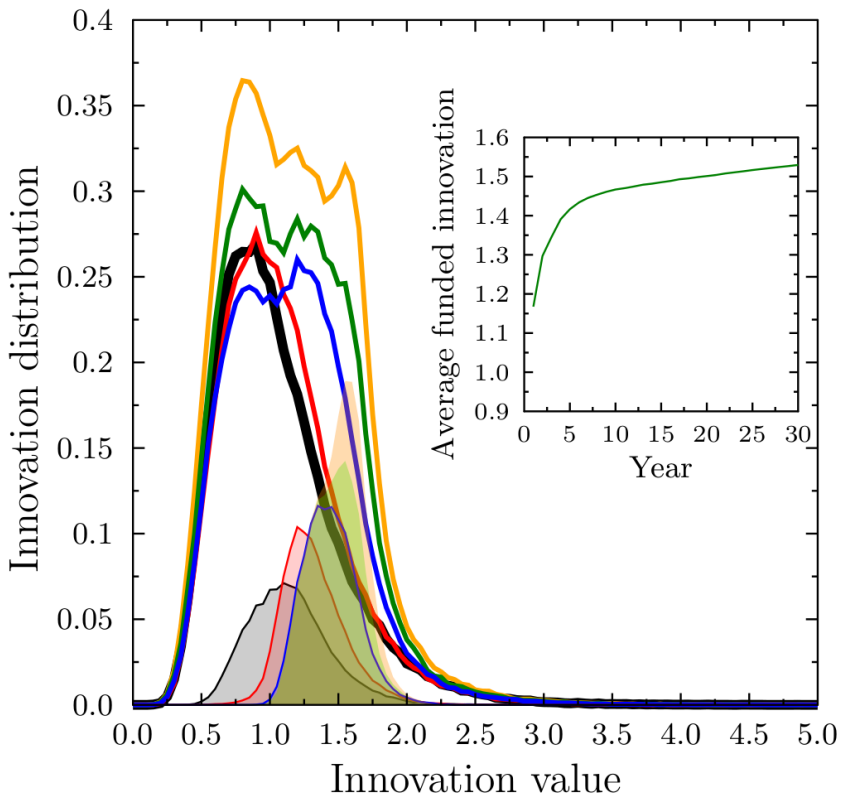


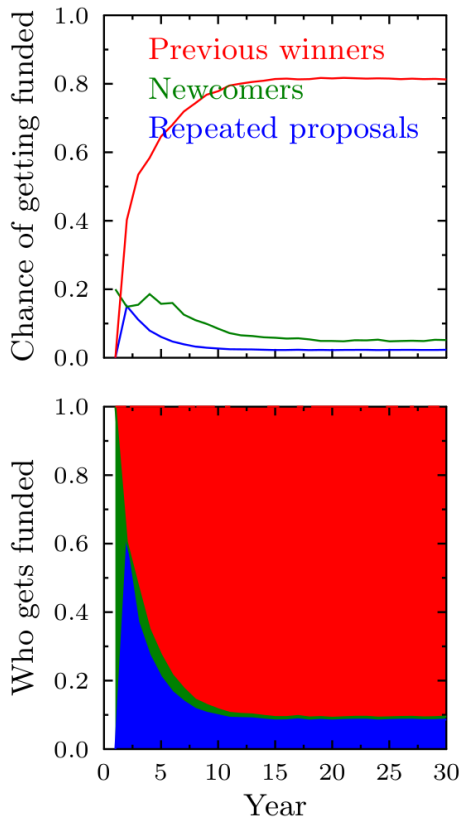
Model results in various circumstances

Bonus for previous succeses (0.1 per evaluation). **Re-evaluation of controversial proposals.**

Low tolerance $\sigma_T = 0.1$. Noise ± 0.3.

Repeated submissions use more of the current 'newspeak'.





Summary

- Unless the reviewers are very open-minded, peer review may indeed favor regression towards mediocrity.

Summary

- Unless the reviewers are very open-minded, peer review may indeed favor regression towards mediocrity.
- Even a relatively weak preference for the current ‘winners’ may lead to disproportionate advantages and biasing the selection process against newcomers .

Summary

- Unless the reviewers are very open-minded, peer review may indeed favor regression towards mediocrity.
- Even a relatively weak preference for the current ‘winners’ may lead to disproportionate advantages and biasing the selection process against newcomers .
- Re-evaluation of controversial proposals by a special, broadminded panel definitely improves the innovation value, but discriminates against newcomers.

Summary

- Unless the reviewers are very open-minded, peer review may indeed favor regression towards mediocrity.
- Even a relatively weak preference for the current ‘winners’ may lead to disproportionate advantages and biasing the selection process against newcomers .
- Re-evaluation of controversial proposals by a special, broadminded panel definitely improves the innovation value, but discriminates against newcomers.
- Special, separate funding scheme for the newcomers is therefore needed.

Summary

- Unless the reviewers are very open-minded, peer review may indeed favor regression towards mediocrity.
- Even a relatively weak preference for the current 'winners' may lead to disproportionate advantages and biasing the selection process against newcomers .
- Re-evaluation of controversial proposals by a special, broadminded panel definitely improves the innovation value, but discriminates against newcomers.
- Special, separate funding scheme for the newcomers is therefore needed.
- *What we did not cover was: individual learning and improvement, systemic biases, fads and fashions, and the top-down driven, politically determined, 'big science' programmes*.

Final quote

Most attempts at innovation, by definition, must fail. Otherwise, they are not truly innovative or exploring the unknown. However, value comes from that small proportion of activities that are able to make significant breakthroughs, as well as from identifying what can be learned from failures.
I have spoken with officials with research funding programmes in the European Commission and in Australia who have acknowledged that despite the brief for their programmes, they are not very innovative.
Instead, they are forced to fund mainly safe projects, for fear of the consequences of failure.
B Perrin, How to – and how not to – evaluate innovation, 2002

Parting question

If we want to explore the unknown, to aim for true innovations, we must accept the risk of failure. This applies – in particular – to research.

Parting question

If we want to explore the unknown, to aim for true innovations, we must accept the risk of failure. This applies – in particular – to research.
The rule of thumb is that 90% of truly audacious efforts end in failure, but the remaining 10% pay off the costs and generate true growth.

Parting question

If we want to explore the unknown, to aim for true innovations, we must accept the risk of failure. This applies – in particular – to research.
The rule of thumb is that 90% of truly audacious efforts end in failure, but the remaining 10% pay off the costs and generate true growth.
Then let me ask the question: do you know any funding agency that BOASTS about the fact that 90% of the research they funded ended in failure?
Because this would mean that they really fund innovation...