Productivisme, productivité et inégalités : des discours aux données empiriques

Enka Blanchard

Centre National de la Recherche Scientifique,
Laboratoire d'Automatique, Mécanique et Informatique Industrielles et Humaines, UMR CNRS 8201, UPHF, CNRS,
Chaire Intelligence Spatiale, UPHF,
Centre Internet et Société, UPR 2000
https://koliaza.com

Joint work with **Zacharie Boubli**, Conservatoire National des Arts et Métiers

Séminaire d'éthique de l'IRMAR

Université de Rennes 1, 17 mai 2024

Definitions (in this context)

To avoid semantic issues, during this talk:

- Productivity: the total production over an extended period
- Productivism: the belief that the goal (of management/public policy) should be to maximise productivity
- Elitism: the belief that objectives (e.g., productivity) are best attained by using unequal distributions of resources

Ideologies are hard to study empirically, so we use a methodology based on immanent critique: Does the policy achieve its stated objectives?

Leo Szilard (1898-1964) on slowing down science

Writing in 1948, before the creation of the NSF (edited for brevity):

> *You could set up a Foundation, with an annual endowment of thirty million dollars. Research workers in need of funds could apply for grants, if they could make out a convincing case; Have ten committees, each composed of twelve scientists, appointed to pass on these applications. Take the most active scientists out of the laboratory and make them members of these committees. And the very best in the field should be appointed as Chairmen [...].*
>
> *The best scientists would be removed from the laboratories and kept busy on committees passing on applications for funds. Secondly, the scientific workers in need of funds will concentrate on problems which are considered promising and are pretty certain to lead to publishable results. [...] Science will become something like a parlor game. Some things will be considered interesting, others will not. There will be fashions. Those who follow the fashion will get grants. Those who won't, will not, and pretty soon they will learn to follow the fashion too.*

Pillars of productivist policies

Assuming productivity is the goal, does public policy achieve its objective?

Current policies seem to rest on three hypotheses:

- one can measure with sufficient accuracy research outputs and productivity (both at the collective and individual levels);
- one can increase the productivity through public policy, using both selection and incentives;
- the cost and adverse effects of these policies are compensated by the productivity gain.

Measurability

Measurability

Three main questions:

- What research outputs do we measure?
- On what temporality (years/careers) and at what scale (individual/team/country)?
- Through which methods: qualitative (peer jury) or quantitative (bubliometry and scientometry)?

Hypothesis zero is that future performance is predictable using past performance.

Types of research outputs

Three main types:

- Teaching, popularisation, supervising, etc.
- Publications of various kinds and participation in peer-reviewed conferences
- economic aspects, both through the industry (patents, startups...) and through ability to obtain funding (ANR, ERC...)

The very presence of the third (especially the ability to obtain funding) shows a shift between objective and indicator.

Bibliometric woes

Bibliometry has been decried for a long time (San Francisco declaration, ARRA).

It is still defended as the only solution against arbitrary choices made by juries.

Italy until 2016: recruitment requires having bibliometric indicator above the median (among people already hired).

Italy now: bibliometric requirements checked by a algorithm whose details are not public.

Garbage everywhere ?

Do metrics give information, or is it "Garbage in, garbage out"?

All metrics fundamentally rely on juries. And Condorcet's jury theorem only works if jurors' opinions are independent.

Are we trading visible biases for hidden biases?

Reviewing issues

Many example of issues:

- NeurIPS 2014: two reviewing committees agree on 51% of accepted papers (random baseline: 23.5%), medium correlation between reviews. No correlation between reviewing score and eventual citations.
- Dougherty and Horne, 2022: in cognitive science, journal impact factor is poorly or even negatively correlated with many statistical indicators and replicability
- Baxt et al., 1998: in medicine, a test of 262 reviewers show that they let through a majority of errors on a test article
- Pier et al., 2018: in medicine, no statistical correlation between reviewers for project evaluations

Cumulative advantages

Matthew effect : initial advantage increases.

For funding, looking at early career funding applications in the Netherlands: no difference with grade among funding recipients, same among rejected candidates. Very large difference between people with similar scores at the threshold.

For citations: retracted papers continue getting cited massively. Citation metrics hugely dependent on journal (even for identical articles).
To publish in Nature: find a co-author who published in Nature.

Goodhart's law

When a measure becomes a target, it ceases to be a good measure.

Good example in effect of Shanghai rankings.

Not only is productivity hard to measure, it gets harder. It is impossible to compare across time because of evolving dynamics.

It is also impossible to compare across fields: median citations rates vary by a factor >30.

Conjecture: an interdisciplinary troll could Dutch-book the system and be the best researcher according to most metrics without being Camille Noûs...

Policies

Impact per dollar

Is the selection system selective?

Study on NIH funding: 1.7M$ in funding gives a 7% increase in publications.
Main hypothesis: multiple parallel funding systems, leading to no selectivity despite high cost

Québec study on 12k researchers: huge productivity increase when going from 0 to 20-50k$ per year then limited marginal increase, then decrease (around 300-400k$).

Prestigious universities have higher success on all funding and publication indicators.
Mid-range universities have more publications (65%) and citations (35%) per dollar.

Incentives

Incentives to publish in prestigious venues can be strong (up to 20 years of salary in countries like China, South Korea, Türkiye).

Switching to extrinsic motivation can be harmful to good work. Study on submissions to Science from 2000 to 2009: financial incentives have a very strong impact on submission rate, and non-statistical negative impact on total accepted papers.

Also creates more opportunities for fraud.

Efficiency

Ideological aspects

No reflexivity: influenced by New Public Management, evolution towards data-driven policies to optimise performance indicators. No checks on whether the change actually increases performance (when counting the overhead).

NHMRC in Australia: on average 35 days of work for a funding application (just for the researchers), 14% of the application's average budget.

In some countries, >40% of working hours are spent on funding applications and funding management. 24% in France according to CNESER.

Gross and Bergstrom, 2019: some funding models have negative expected returns (when including evaluation costs).

Most success rates 10-20%. Dangerous cycle: lower success rate increases randomness, increases incentive to spam, reduces reviewing quality, increases randomness.

Let them research cake

We might have enough funding in practice. ANR budget: >10k€ per researcher (including doctoral students).

Study on Uk and Netherlands: keeping doctoral recruitment constant, total funding is >32k€ per year.

Bet : most French researchers would accept 5-10k€/year (plus doctoral students) in exchange for never applying for more funds

Changing is hard

Changing evaluation methods is hard. Study on Australian medical funding: lowering page limits for funding applications increases the time spent.

We are reticent to study our own practices (because it would be unfair, like a placebo study). Pierre Azoulay:

> *I encounter a lot of resistance. Wouldn't this be gambling with scientists' careers? How can we measure success — by counting publications and citations, looking at the students trained as a by-product of these grants, or using other metrics? Won't this work shift scarce funding away from actual scientific investigations?*
>
> *These criticisms are without merit. The current system already gambles with scientific careers, just in a haphazard way. [...]*
>
> *Experimenting on ourselves may well lay bare some shortcomings of the scientific community and expose us to criticisms from politicians, who are always looking for excuses to cut science funding. But the only alternative to such controlled experimentation is the gradual stultification of our most cherished scientific institutions.*

Conjecture: dismantling competitive funding could increase research "productivity" by ≈ 2

Alternatives

Multiple options seem to be better:

- Dismantle all competitive funding, fund researchers as much as they want (up to an equal share)
- Use a lottery for funding big projects (statistically equivalent to actual policy but less costly). Already on trial in Aotearoa New Zealand
- Use a recommendation system where each node redirects funding, which limits the impact of cartels (Bollen et al., 2017)

Summary

Current research policies on funding/recruitment:

- do not reliably select the highest performers;
- have large negative externalities;
- are very costly.

Why ? Ideology seems to bear some responsibility (otherwise, it would be a conscious effort to destroy the system). Sometimes explicitly so (e.g., architect of New Public Management in the USA).

Thank you for your attention

To go further:

- *Recherche et dogmatisme : de l'improductivité du productivisme*, joint work with Zacharie Boubli, published online (open access) in Questions de Communication.
  `https://journals.openedition.org/questionsdecommunication/29994`
- Conference and debate with Bernard Friot on research ethics at INRIA Rennes (2023).
  `https://atelier-ethique-inria2023.gitlab.io/`

Time for questions.