Open Science, Closed Peer Review?

A Case Study of Economics Researchers

Adam Gill

Uppsala University
Swedish Reproducibility Network
Lab x Lab

November 10, 2023

Questions for today

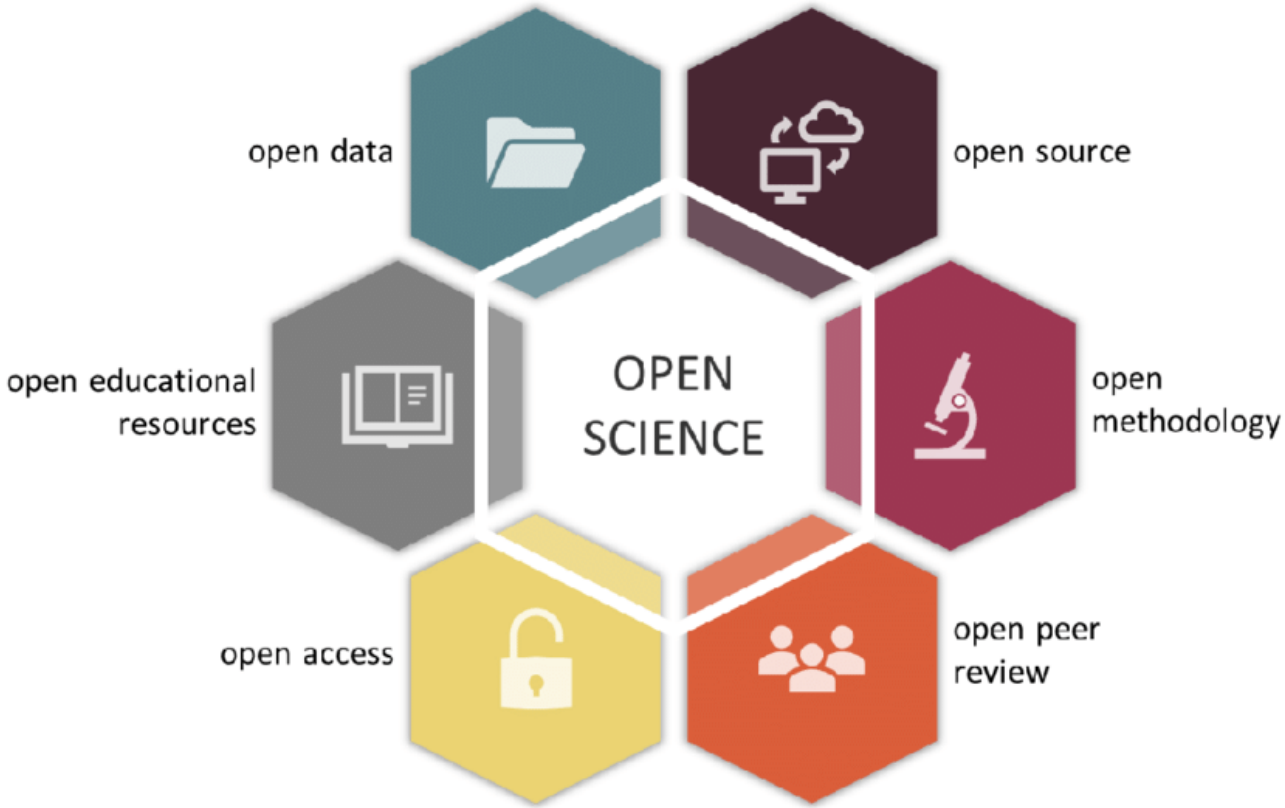
- The main questions related to this talk:

1. What is open peer review?
2. How has it been implemented?
3. What aspects have not been implemented and why?
4. Do we even want more open peer review?

Introduction

- Open Science movement has been growing in recent years, both in general and (more slowly) in economics
  - Increased use of pre-analysis plans
  - Greater reproducibility policies at journals
- Peer review, however, is often opaque despite its importance to research
  - Open Peer Review: Catch-all for policies that increase the transparency of and access to peer review process and data
- But these policies come with real costs!
- Use survey responses from (mostly) economics researchers and a review of the literature to shine some light on this

Pillars of Open Science



Source: Falster et al., 2019

Defining open peer review

- *Open Peer Review* is an umbrella term for all peer-review-related policies that follow the general principles of open science
- There are generally three types of policies that are associated with open peer review
  1. Policies that modify the identifiability of authors, referees, and editors
  2. Policies that facilitate the disclosure and exchange of peer-review documents, metadata, and other information
  3. Policies that make participation more accessible and widespread
- Across disciplines, open peer review practices have been adopted to different degrees, but it has been adopted much less quickly than other open science practices

The survey

- Survey about researcher's experience with the peer review system, opinions on the current performance, and attitudes towards proposals
- Who: Economics researchers who have written & received at least one peer review report over the previous 2 years
- When: July 2020 - January 2021
- Total of 1,459 fully completed responses
- Over-sampled experimental economists and European researchers; under-sampled female researchers
- Results are based on the experimental/behavioral economics subset (N = 802), but they are qualitatively similar to the full sample

The Context of Economics

- Economics seems similar to other social science disciplines when comparing a small sample of top journals (but different from general science and medicine)

Compared to psychology and political science journals:

- Similar number of publications, impact factor, h-index, and some administrative aspects
- Economics has no double-blind or suggesting/opposing of referees, unlike political science or psychology
- None of these disciplines have public referee identification or public reports

Comparison between disciplines

Table: Comparison of top journals across disciplines

General	*Economics*	*General Science*	*Medicine*	*Political Science*	*Psychology*
Professional Editorial Board	0 of 5	2 of 3	1 of 3	0 of 3	0 of 3
Articles Published (2019-2021)	241	6,259	1,718	295	281
Impact Factor (two-year)	9	16	23	5	14
h-index	241	1,151	908	174	362
Average Days from Submission to Acceptance	752	–	–	437	437
Open Peer Review Policies					
Double-Blind Refereeing	0 of 5	1 of 3	0 of 3	3 of 3	3 of 3
Public Referee Identity	0 of 5	2 of 3	0 of 3	0 of 3	0 of 3
Public Referee Reports	0 of 5	1 of 3	0 of 3	0 of 3	0 of 3
Authors Can Suggest/Oppose Referees	0 of 5	3 of 3	1 of 3	3 of 3	2 of 3

Identifiability

- Three main types of peer review systems with respect to identifiability of stakeholders (COPE, 2019):
  - Double-blind: Authors and reviewers anonymous to each other
  - Single-blind: Reviewers anonymous to authors but not vice versa
  - Open (no-blind): No anonymity in the process
- There could also be a triple-blind system - editors are also blinded to authors

Anonymity of authors

- Why make authors anonymous?
  - Protect against identity-based bias
  - Authors may feel that they get a "fair trial"
- Arguments against:
  - Evidence of its effectiveness at reducing bias is mixed
    - Likely field specific
  - Disciplines where pre-prints are common and widely circulated, anonymity becomes difficult to maintain (Rastogi et al., 2022)
  - Similarly, increased pressure for pre-analysis plans or registered reports making author anonymity nearly impossible

Anonymity of referees

- Why make referees anonymous?
  - May protect referees from retaliation, favor-swapping, or social pressure (Tennant et al., 2017)
  - Removing it may disincentivize participation (Bianchi and Squazzoni, 2022)
- Arguments against:
  - Limited accountability for what they write (Tennant et al., 2017)
  - Could write reports that are mean, untrue, or aggressive without consequences
    - May affect underrepresented groups more often

Attitudes towards identifiability policies

- Author Identifiability:
  - Mixed opinions: 39.3% had favorable views while 46.4% had unfavorable views towards a double-blind system
- Referee Identifiability:
  - Respondents are generally negative: 60% said lifting the anonymity of senior referees would not be very useful
  - However, it could be needed: respondents sometimes received low quality reports with the low quality coming the most from "inaccurate statements" (75.6%), but also a non-negligible amount coming from "aggressive tone" and/or "personal insults" (41.9%)
    - Greater transparency could increase accountability and reduce this
- Editor Identifiability:
  - More mixed: 40% said it would not be very useful while 42% said it would be useful

How has it been implemented?

- Top economics journals all use single-blind, but other economics-related journals (e.g. *Management Science*) and top psychology journals (e.g. *Psychological Bulletin*) use double-blind
- Triple-blind is used in some top philosophy journals (e.g. *Philosophical Review*)
- Some journals (e.g. *Nature*) have an opt-in review system

Publication of peer review documents

- Could publish referee reports anonymously
  - Pros: Could incentivize better reports without removing anonymity ; Could get better data on how marginal improvements affect papers
  - Cons: May not actually affect referee behavior; May disincentivize participation as referees
  - *Note:* Anonymity may be hard to keep in some cases due to writing style, formatting, etc.
- Referees sign their reports and the entire review history is disclosed
  - Pros: Strongly promote accountability
  - Cons: Likely reduces participation as referees; no discernible evidence of report quality improvement (van Rooyen et al., 2010)

Gradient of open peer review

- **Full transparency**: Referee reports are published with names
- **High transparency**: Referee names and reports are published but not connected
- **Moderate-High transparency**: Reports are published with manuscripts, but not referee names
- **Moderate transparency**: Referee names are published with manuscripts, but not reports
- **Moderate-Low transparency**: Opt-in system for referee revealing their names
- **Low transparency**: Referee names are revealed at a journal-level (maybe with number of reports written or other metadata)
- **No transparency**: No information about referee reports or names is revealed

Attitudes towards publication of peer review documents

- Our respondents were somewhat favorable (55%) to publishing referee reports in an anonmyized way
  - In a cross-disciplinary survey, respondents were more favorable (Ross-Hellauer et al., 2017)
  - Despite being favorable, only about 38% expect this policy to be useful for peer review
- Our respondents were skeptical about a fully open peer review system both for a system that applies to all referees and one that only applies to senior referees
  - This mirrors patterns among other social science researchers (Publishing Research Consortium, 2016)

How has it been implemented?

- *Frontiers* publishes referee names with manuscripts conditional on acceptance
  - Could also have an opt-in system like e.g. *Nature* or *Advances in Methods and Practices in Psychological Science* (*AMPPS*)
- *Nature* and *AMPPS* also publish the anonymized reports with the manuscripts (unless referees opt-in to include their names as well)
- Full transparency, where both reports and names are published together, is only embraced by a handful of journals (e.g. *F1000Research* and *The BMJ*)
- Economics journals primarily follow the "standard" peer review model (no publishing of reports or referee names)

Publication of peer review data

- Increased sharing of peer review metadata, while less transparent, could also be useful
- Journals could periodically post detailed delay and rejection statistics
  - Pros: Create public accountability; allow authors to make more informed decisions
  - Cons: Journals could make changes targeted at statistics, not improving the process; authors may be discouraged or disincentivized from submitting at certain journals (e.g. high rejection rate)
- Publish summaries of how manuscripts changed from original submission to published version
  - Pros: Adds accountability; insights into the benefits/costs of revision process
  - Cons: May discourage certain types of revision requests

How has it been implemented?

- While many journals offer some statistics about their refereeing and publication, it is often limited, inconsistent, and non-comparable across journals
  - Few journals seem to publish consistent lists of researchers who referee or statistics on turnaround times
- While, the amount of statistics and number of journals offering information is increasing, more can still be done
- E.g. *F1000Research* is one of the few journals that publishes summaries of how manuscripts changed from submission to publication

Open channels of communication

- Increasing the frequency and ease of communicating between different participants
- Types of Open Communication:
  - Create an official channel for authors to appeal a decision, even without guarantee of response
  - Additional channels of communication at various levels (such as after revise and resubmit, R&R) in order to clarify comments/understand which comments are important
  - Fully interactive review phase where authors and reviewers can interact (maybe through editor as intermediary) in real time or at least through repeated exchange

Arguments about open communication

- Arguments for:
  - Could speed up the review process
  - Could give authors a stronger voice
- Arguments against:
  - Could place more demands on already over-loaded editors and referees
  - Could (unintentionally) benefit certain groups over others

Attitudes towards open channels of communication

- $>$ 60% of respondents would like an official channel for submitting an appeal about an editor decision
  - Only mildly weaker support among editors and active reviewers, despite extra work involved
- Additional communication channels after R&R could help with unrealistic/inconsistent demands, which was mentioned by 46% of our respondents
- Respondents in a different survey were quite supportive of a more interactive review phase with 68% being in favor of this policy (Ross-Hellauer et al., 2017)

How has it been implemented?

- Many journals have unofficial appeal processes and some (e.g. *American Economic Review*) state official appeal processes, but do not discuss how the process works
    - More working can be done on formalizing this process
- *Frontiers* is one of the few journals with an interactive review process, where a real-time discussion forum follows an independent review phase

Who should be the gatekeepers?

- Refereeing is often dominated by a small group of researchers
  - Mean (median) number of reports written is 10.2 (8) per person per year
  - The top 10% most active referees wrote 20 or more annually
- Who should be responsible for refereeing?
  - Relying on a handful of trusted, experienced referees could create dependable signals and may be more efficient
  - But it can create excessive workloads on some researchers, may stifle innovative projects, and puts the publication power in the hands of only a few people which can affect other researchers careers

Changes to the referee pool

- Could increase the number of referees
  - Pros: Improve inter-reviewer correlations; improve the average informativeness of signals; more participation in the process
  - Cons: Increases burden of system; may only provide marginal benefit
- Could have "shadow reviewers" or "specialized reviewers" as well
- Use the "wisdom of the crowd"
  - Let all researchers (or even all of society) determine what is good and useful research
  - Public "ranking" system (e.g. Yelp or Rotten Tomatoes for research) - Could weight by expertise
  - Pros: More involvement and more signals; less power in the hands of a few
  - Cons: Need to reduce the extensiveness of the assessment; Could exacerbate biases
- Could combine traditional peer review with "wisdom of the crowd"

Next steps to take

- More discussion (like today!) about if and how much open peer review we want
- Make metadata on peer review more available, transparent, and consistent
- Experiment with different policies to see what works
- Create a cross-disciplinary panel to discuss best practices and (maybe) make policies more uniform across disciplines

Swedish Reproducibility Network (SweRN)

- The Swedish Reproducibility Network (SweRN) is a peer-led network of researchers with the goal to make research results more reliable
  - Promote and ensure rigorous and reproducible research practices in Sweden
- Our goal is to involve researchers in Sweden from all disciplines

Lab^2 (Lab x Lab)

- Lab^2 is a lab for experimentation in science that aims to carefully document the research process and learn how to improve it
- The main mission is to reduce inefficiencies of production and inequalities of access to information by centralizing existing knowledge and producing new evidence via coordinated efforts.

Ways to help

- Contribute metadata
  - Background behind your published work
  - New insights from unpublished work
  - Any "building blocks" of your research projects
- Share your views
  - Making predictions
  - Providing direct feedback
- Join large studies
  - Participate in systematic replications
  - Take part in multi-design or analyst studies
  - Join randomized trials

Thank you for listening!

References I

Bianchi, F., & Squazzoni, F. (2022).Can transparency undermine peer review? A simulation model of scientist behavior under open peer review. *Science and Public Policy*, *49*(5), 791–800. https://doi.org/10.1093/scipol/scac027

COPE. (2019).Ethical guidelines for peer reviewers [COPE Guidelines are formal COPE policy and are intended to advise editors and publishers on expected publication ethics practice]. https://doi.org/10.24318/cope.2019.1.9

Falster, D., Maitner, B. S., Salguero-Gómez, R., & Vandvik, V. (2019).The open traits network: Using open science principles to accelerate trait-based science across the tree of life.

Publishing Research Consortium. (2016, May). *Peer review survey 2015* (tech. rep.).

References II

Rastogi, C., Stelmakh, I., Shen, X., Meila, M., Echenique, F., Chawla, S., & Shah, N. B. (2022).To arxiv or not to arxiv: A study quantifying pros and cons of posting preprints online. *arXiv preprint arXiv:2203.17259*.

Ross-Hellauer, T., Deppe, A., & Schmidt, B. (2017).Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers. *PloS one*, *12*(12), e0189311.

Tennant, J., Dugan, J., Graziotin, D., Jacques, D., Waldner, F., Mietchen, D., Elkhatib, Y., Collister, L., Pikas, C., Crick, T., Masuzzo, P., Caravaggi, A., Berg, D., Niemeyer, K., Ross-Hellauer, T., Mannheimer, S., Rigling, L., Katz, D. S., Tzovaras, B., & Colomb, J. (2017).A multi-disciplinary perspective on emergent and future innovations in peer review. *F1000Research*, *6*, 1151. https://doi.org/10.12688/f1000research.12037.3

References III

van Rooyen, S., Delamothe, T., & Evans, S. J. (2010).Effect on peer review of telling reviewers that their signed reviews might be posted on the web: Randomised controlled trial. *BMJ*, *341*(nov16 2), c5729–c5729. https://doi.org/10.1136/bmj.c5729