

## Rethinking Statistical Publication

David Banks, Duke University

I believe most people agree that there is no need for physical journals in the future. Our profession does not access information in that way anymore. Tying publication to a physical journal model has pernicious consequences:

1. *JASA* will not accept a paper longer than 35 pages, but sometimes a paper needs to be long (e.g., a review paper may need to be lengthy).
2. There is always an unnecessary email exchange with the publisher about grayscale and color. We should all be using color routinely---it is better.
3. Authors cannot include video in their papers. But for spatial processes or dynamic networks or many other topics, video is more effective than text or static images.
4. It isn't possible to link code to the paper. But readers might want to modify the code, e.g., in order to explore other parameter settings for simulations.
5. We cannot seamlessly link to data from inside the paper.

Physical journals are expensive to produce, bad for the environment, and inefficient. Most researchers only read articles online. Moreover, few people read more than one article in an issue of *JASA* or any other ASA journal. Individuals who prefer physical copies will typically print a single article rather than read from a hardcopy journal.

The economic model for publication is outdated. A researcher works to write a good paper, then signs a copyright transfer agreement so the publisher can own the paper for free and put it behind a paywall in order to make money by selling it to our friends. That model made sense when the only way to share one's ideas with the professional community was through hardcopy distribution. But that constraint disappeared at least 20 years ago. We can post papers online.

Our review system is not sustainable. Generative AI and the increased number of submissions are already straining the system.

1. *JASA* is three times the size it was just 15 years ago, and we have seen an increase in the number of new journals (for which I am partly to blame).
2. If a paper is rejected at a top journal, it gets sent to a lesser one, and then climbs down the prestige ladder until some journal accepts it or the author gives up.

Each time, it gets assigned to two reviewers, which is extravagantly wasteful of that limited resource.

3. Senior faculty rarely review, and junior faculty should focus on their research instead of reviewing lots of papers.
4. The quality of reviews is uneven---we all know of cases in which a review added significant value, but there are many cases in which it does not, and sometimes the review damages the paper.
5. The review process guarantees about a year's delay in publication.

It is long past time to review the costs and benefits of our review processes. We handle papers just as we did 80 years ago, even though almost every aspect of how we acquire, use and evaluate information has changed.

I have spoken with many people about ways to improve our system. Some want to pay reviewers, some want to move to computer-science-style refereed conference proceedings, and there are other models. The one I recommend is to post our papers on arXiv.

The arXiv system is essentially free, thanks to Jim Simons. Papers are not reviewed but they *are* moderated---if someone misbehaves, their paper will be taken down. (I was a moderator for almost a decade, and it works well---readers alert a moderator if a paper denies the Holocaust or is unprofessionally rude.)

The arXiv system has many good features:

1. One can incorporate color, video and code.
2. The author retains ownership.
3. It contains living documents---the author can update a paper or fix an error and post the improved version on top of the older one (so one should always cite the time stamp on which an arXiv paper was accessed).

The one thing arXiv does not provide is a quality flag. In the current system, the prestige of a journal provides a noisy signal about the quality of the paper. That signal can be problematic---if a paper appears in a prestigious journal, readers may trust it more than they should.

It is easy to add a quality flag to arXiv. One could write a wrapper that uses a Google page rank algorithm to assess quality. If a highly cited author cites an arXiv paper, that indicates value. If an undergraduate cites that paper, not so much. One could tweak

the algorithm to also take account of the number of highly cited authors citing the paper, making the system less gameable than the one we currently have.

There are many variations on this theme. For example, an author could post a paper on arXiv and then ask someone designated by the ASA or IMS as an editor to consider accepting that paper as, say, a *JASA* paper. The editor could launch a standard review process and ultimately decide whether to give that paper a *JASA* certification. But publishers are irrelevant to that system.

There are other details to consider, but I hope the main message is clear. I am willing to live with the present system if our professional societies choose that path, but it should be decision and not simply a matter of habit.