Testing for Citation Bias in Conference Peer Review

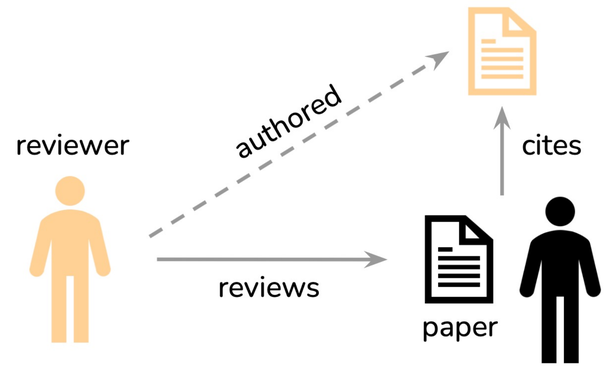
Ryan Liu advisor: Nihar Shah collaborators: Ivan Stelmakh, Charvi Rastogi

Carnegie Mellon University, School of Computer Science

Research Problem

Does the **citation of a reviewer's work** in a submission cause the reviewer to be positively biased, causing a shift in evlauation that goes beyond the genuine change in the submission's scientific merit?

We measure the effect that comes **on top** of the objective improvement achieved by having a stronger bibliogrpahy.



Background

- Citations have a large influence on researchers' careers, from Google Scholar first impressions to hiring and promotion decisions[1][2]
- Anecdotes suggest that citing papers of likely reviewers is effective in increasing the chances of acceptance[3]
- Unethical reviewer behaviors similar in power dynamic occur in both extreme examples[4][5] and general studies[6][7]
- To match reviewers, authors are also likely to be proactive in these scenarios[6]
- Two past works[8][9] touch on cited reviewers being favorably biased, but do not account for important confounders

Methods

- We perform two observational experiments in conferences ICML 2020 and EC 2021
- To increase statistical power in EC, we created a novel assignment algorithm that jointly optimizes #cited reviewers and traditional similarity
- Citation relationship is an assisting indicator for reviewer-paper expertise

	ICML 2020	EC 2021
# Reviewers	3,064	154
# Submissions	4,991	496
# Submissions with > 0 cited reviewers	1,513	287
Fraction of submissions with > 0 cited reviewers	30%	58%

Analysis: confounders

- Genuinely Missing Citations
  Authors of missing important works will more likely notice
- Paper Quality
  Higher quality papers may have more citations
  Papers better fit for the conference cite more reviewers
- Reviewer Expertise
  Cited reviewers may have more expertise
  Expertise causes homophily and strategic behavior
- Reviewer Preference
  Assignment jointly optimizes preference and citation
- Reviewer Seniority
  Senior reviewers are more widely cited
  Senior reviewers more lenient

Analysis: techniques

- Data Filtering: Reviewers label if papers were missing important citations, inspected and removed true positives
- Parametric Inference: Using confounding variables, we model:

$$\texttt{score} \sim \alpha_0 + \alpha_1 \cdot \texttt{quality} + \alpha_2 \cdot \texttt{expertise} + \alpha_3 \cdot \texttt{preference} + \alpha_4 \cdot \texttt{seniority} + \alpha^* \cdot \texttt{citation}$$

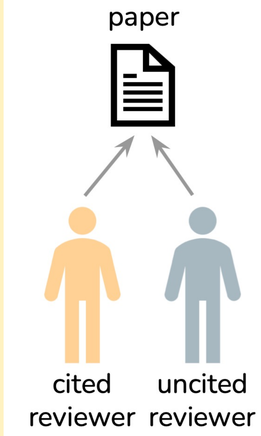
  Comparing pairs of (cited, uncited) reviewers within the same paper, we cancel out the unknown paper quality and test for significance of α^* using weighted linear regression.

$$\texttt{score}_\Delta \sim \alpha_2 \cdot \texttt{expertise}_\Delta + \alpha_3 \cdot \texttt{preference}_\Delta + \alpha_4 \cdot \texttt{seniority}_\Delta + \alpha^*$$

- Nonparametric Inference:

  Strong modeling assumptions are not guaranteed to hold in peer review settings[10], motivating a test that does not assume linearity. For the ICML data, we find (cited reviewer, uncited reviewer, paper) triples, where the reviewers share similar expertise, seniority, and bids.

  We run a permutation test on all these pairs, permuting the citation labels within pairs for 10,000 iterations.



self-reported expertise

self-reported confidence

semantic text analysis

$|S(\text{cited}, \text{paper}) - S(\text{uncited}, \text{paper})| \leq 0.1$

bids are both 3 ("in a pinch"), or ≥ 4 ("willing")

seniority indicator

Results

EC 2021	Parametric	0.23 (p = 0.009)	5 pt scale
ICML 2020	Parametric	0.16 (p = 0.004)	6 pt scale
ICML 2020	Nonparametric	0.42 (p = 0.02)	6 pt scale

- A single-point increase in score from one reviewer contributed to an 11% increase in chance of acceptance

Citations

[1]: Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences, 102(46):16569–16572.
[2]: Fuller, S. (2018). Must academic evaluation be so citation data driven? https://www.universityworldnews.com/post.php?story=20180925094651499
[3]: Kostoff, R. N. (1998). The use and misuse of citation analysis in research evaluation. Scientometrics, 43(1):27–43.
[4]: COPE (2018). Editor and reviewers requiring authors to cite their own work. https://publicationethics.org/case/editor-and-reviewers-requiring-authors-cite-their-own-work
[5]: Van Noorden, R. (2020). Highly cited researcher banned from journal board for citation abuse. Nature, 578:200–201.
[6]: Fong, E. A. and Wilhite, A. W. (2017). Authorship and citation manipulation in academic research. PLoS ONE 12, 12.
[7]: Resnik, D. B., Gutierrez-Ford, C., and Peddada, S. (2008). Perceptions of ethical problems with scientific journal peer review: An exploratory study. PMC, 14(3):305–310.
[8]: Sugimoto, C. R. and Cronin, B. (2013). Citation gamesmanship: Testing for evidence of ego bias in peer review. Scientometrics, 95(3):851–862.
[9]: Beverly, R. and Allman, M. (2013). Findings and implications from data mining the imc review process. SIGCOMM 2013, 43(1):22–29.
[10]: Stelmakh, I., Shah, N., and Singh, A. (2019). On testing for biases in peer review. In NeurIPS.