DOES THE MARKET OF CITATIONS REWARD REPRODUCIBLE WORK?

Booz | Allen | Hamilton

Edward Raff[1,2]

1 Booz Allen Hamilton 2 University of Maryland, Baltimore County

ML Evaluation Standards Workshop at ICLR 2022, 29 April 2022, Virtual

UMBC

Abstract

Citations are one of the primary incentive and reward systems for academic work, and so we desire to know if this incentive rewards reproducible work. Yet to the best of our knowledge, only one work has attempted to look at this combined space, concluding that non-reproducible work is more highly cited. We show that answering this question is more challenging than first proposed, and subtle issues can inhibit a robust conclusion. To make inferences with more robust behavior, we propose a hierarchical Bayesian model that incorporates the citation rate over time, rather than the total number of citations after a fixed amount of time. In doing so we show that, under current evidence the answer is more likely that certain fields of study such as Medicine and Machine Learning (ML) do correlate reproducible works with more citations, but other fields appear to have no relationship. Further, we find that making code available and thoroughly referencing prior works appear to also positively correlate with increased citations.

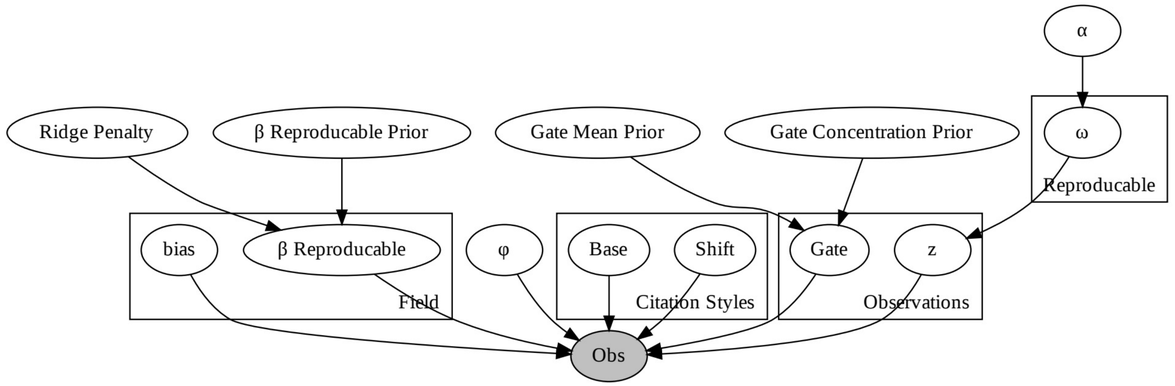
Study Design

- *3 original replication datasets*
- *1 added replication dataset*
- *1 ML specific replication dataset*
- *Prior work sub-selected data and had modeling deficiencies that make hard-to-trust conclusions*
- *New approach is more robust, using all the data indicates no or positive correlation between replication and citations!*
- *Bayesian hierarchical model to account for nuance and limited data*

Hierarchical Model of Citation

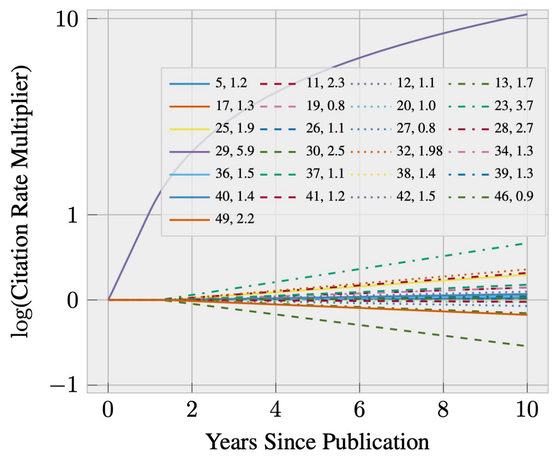
We need to account for differences between fields!

- Each field gets its own bias term for different intrinsic citation rates.
- The citation rate is altered over time (instead of cumulative) by a multiplicative base^(time-shift)
  - This accounts for changes over time
- A single pool of possible “styles” that each paper must select from
- Each population of (non)reproduced papers gets it’s own Dirichlet Process (DP) prior over the pool, so we can determine if there are differences in styles.
- DP hyper-prior from a Beta distribution to enforce a sparsity educing hyper-prior for interpretability reasons.



Matches Prior Bibliometric Results

- Our model learns to shift the base rate change starting 2 years out from publication, matching prior bibliometric results. Strong evidence for our model



Results