Comparing Peer Review Regimes in an Epistemic Landscape

An Agent-Based Model

Carlo Debernardi, Francesco Renzini

carlo.debernardi1@unimi.it,francesco.renzini@unimi.it

UNIVERSITÀ DEGLI STUDI DI MILANO

UiA Universitetet i Agder

BEHAVE

13/09/2022 - SSC2022

Why an ABM on peer-review

- Peer-review is the most common selection process in science
- Awareness of its limits and Open Science movement spurred the adoption of several different regimes
- The differences are difficult to explore due to (Feliciani et al., 2019):
  - *data availability* (confidentiality, especially in case of rejections)
  - difficulties in setting up *experiments*
- Use of ABMs to overcome these difficulties and gain at least counterfactual insights on different peer-review regimes

Previous works

- Thurner and Hanel (2011), Squazzoni and Gandelli (2012, 2013) and Bianchi and Squazzoni (2022)
- How reviewers' behaviors in different regimes could influence the *efficacy* of a field
- Efficacy = ability to select papers with the highest intrinsic quality for publication
- Key findings:
  - Strategic behavior (e.g., indirect reciprocity) reduces efficacy (compared to baseline: Reliable reviewers)
  - Biases in publication are magnified in open peer review regime (status, direct reciprocity)

Limitations and contribution

Fundamental limitation of previous ABMs

Intrinsic quality of manuscripts as proportional (albeit in a noisy way) to the resources the author can access

Our argument

Resources are an enabling factor, but the final quality of a contribution depends on the specific combination of TDM (theory, data, methods etc).

How to model quality

Incorporating all these elements in a model is far from trivial. We draw on contributions from ABMs developed in the field of philosophy of science.

- Weisberg and Muldoon (2009): spatial representation of cognitive dimensions of research TDM, epistemic landscape
- Alexander, Himmelreich, and Thompson (2015): generalization to multi-dimensional space that allows for more complex dynamics
- Sobkowicz (2017): adds steering from funders and sub-discipline topology
- Avin (2019): dynamic epistemic landscape

Our model (motivation)

For these reasons, we developed an ABM with the aim of:

- Comparing the efficacy of different peer review regimes (double and single blind, open; Squazzoni and Bianchi, 2022)
- By using a more complex definition of manuscript quality based on the NK-model (Kauffman & Levin 1987)
- Identifying more *nuanced* and *non-linear* relationships between reviewers' strategies, publication bias and development of a scientific field

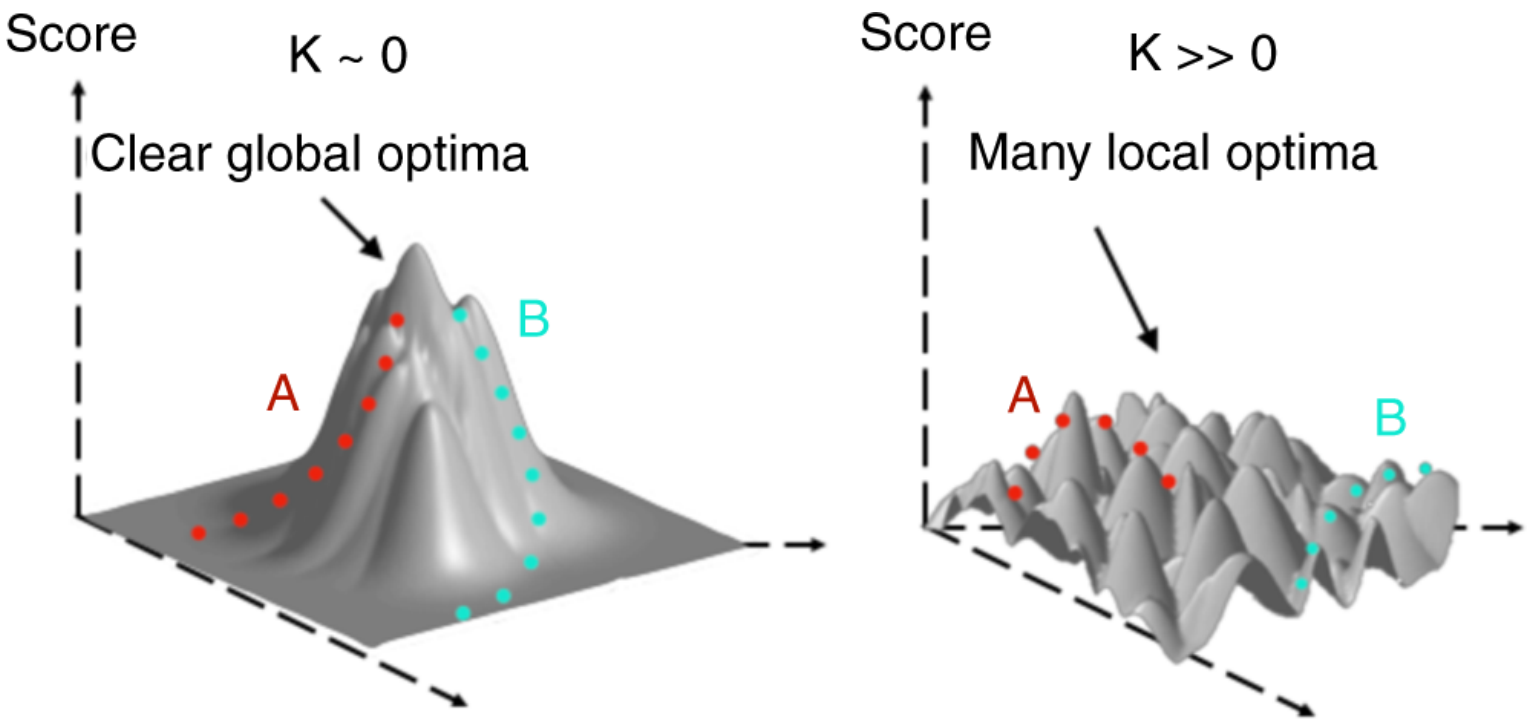
NK-model

Model originally introduced in evolutionary biology (Kauffman & Weinberger 1989), then applied in a handful of research areas like management, organizational studies, ABMs, etc.

Each point in a N-dimensional space is a possible "genome" (TDM-combination) for an agent. Each of the N "genes" contributes to the global fitness (scientific significance) depending on his own value combined with the K-neighbors.

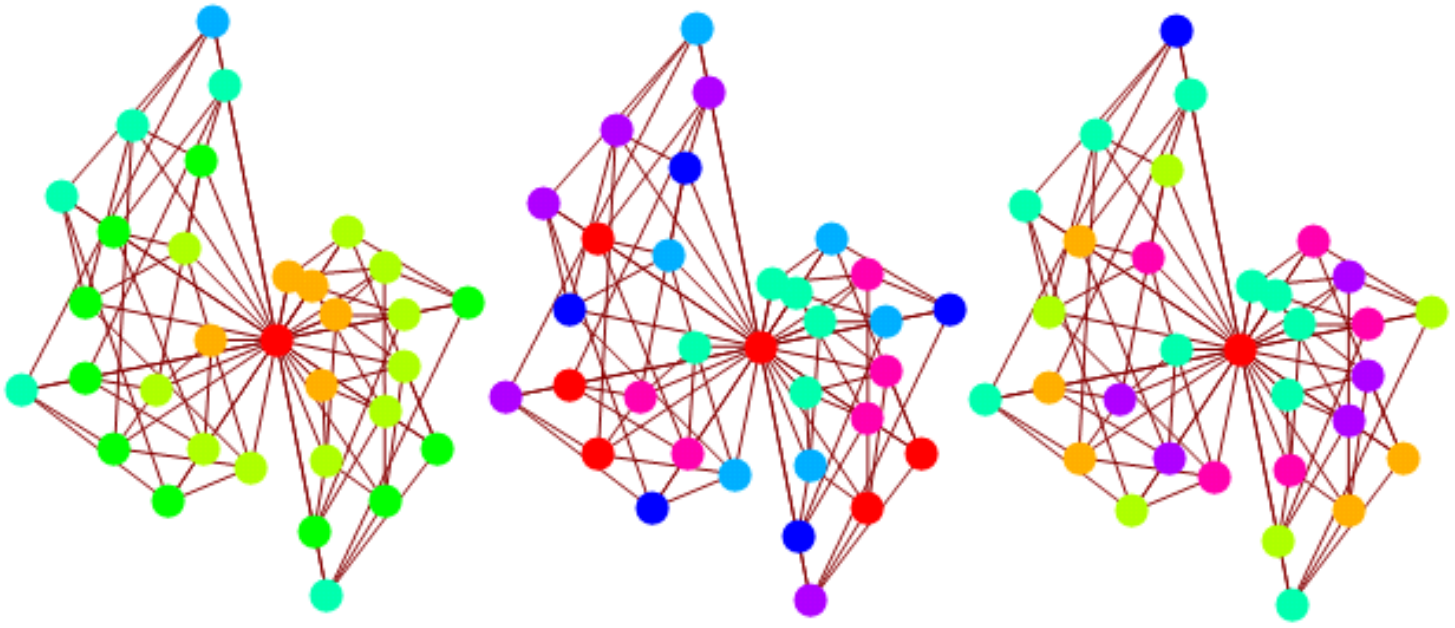
K=0 trivial case of independence. High values of K imply a more rugged landscape, more local optima and a more difficult search process (epistasis).

NK-model



NK fitness landscape

NK-model



N=5, K=[0, 1, 2]

Our model (details)

Algorithm 1 Main Loop of our model

Require: $A > 0$ (number of agents); N; K; d(distance inflation factor); o (over/underrating factor); p (# of published papers); RR (review regime)
Initialize NK, resources, # of agents;
while $t \leq T$ **do**
compute fitness of TDM
Set submission quality $q_m = min\{fitness, resources\}$
Match author-reviewers randomly
Set evaluated quality q_e, according to review regime
Publish top-p papers, according to q_e ranking
Update resources for published authors
Update NK-potential scientific significance and fitness
Move authors (i.e., change randomly TDM)
end while

Preliminary results - Common dynamics

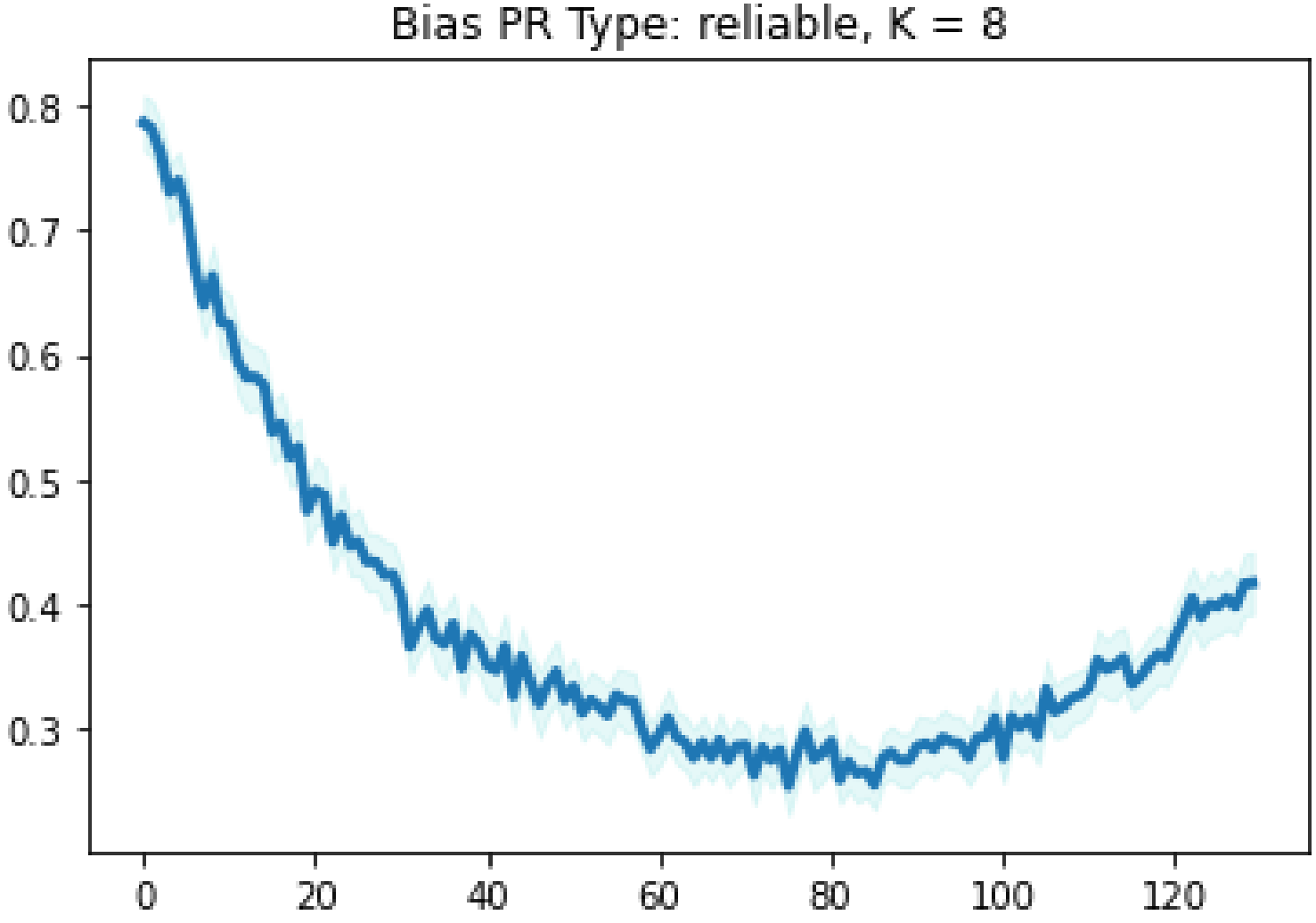
At the beginning resources are uniformly distributed between $[2; 3]$, while fitness is between $[4.5; 6.5]$; resources as a constraint, when everything is yet to explore

At the beginning of a scientific field, q_m is determined by differentials in resources

$p = 10, A = 100, T = 130, d = 1.5, o = 2, K = 8, N = 10$

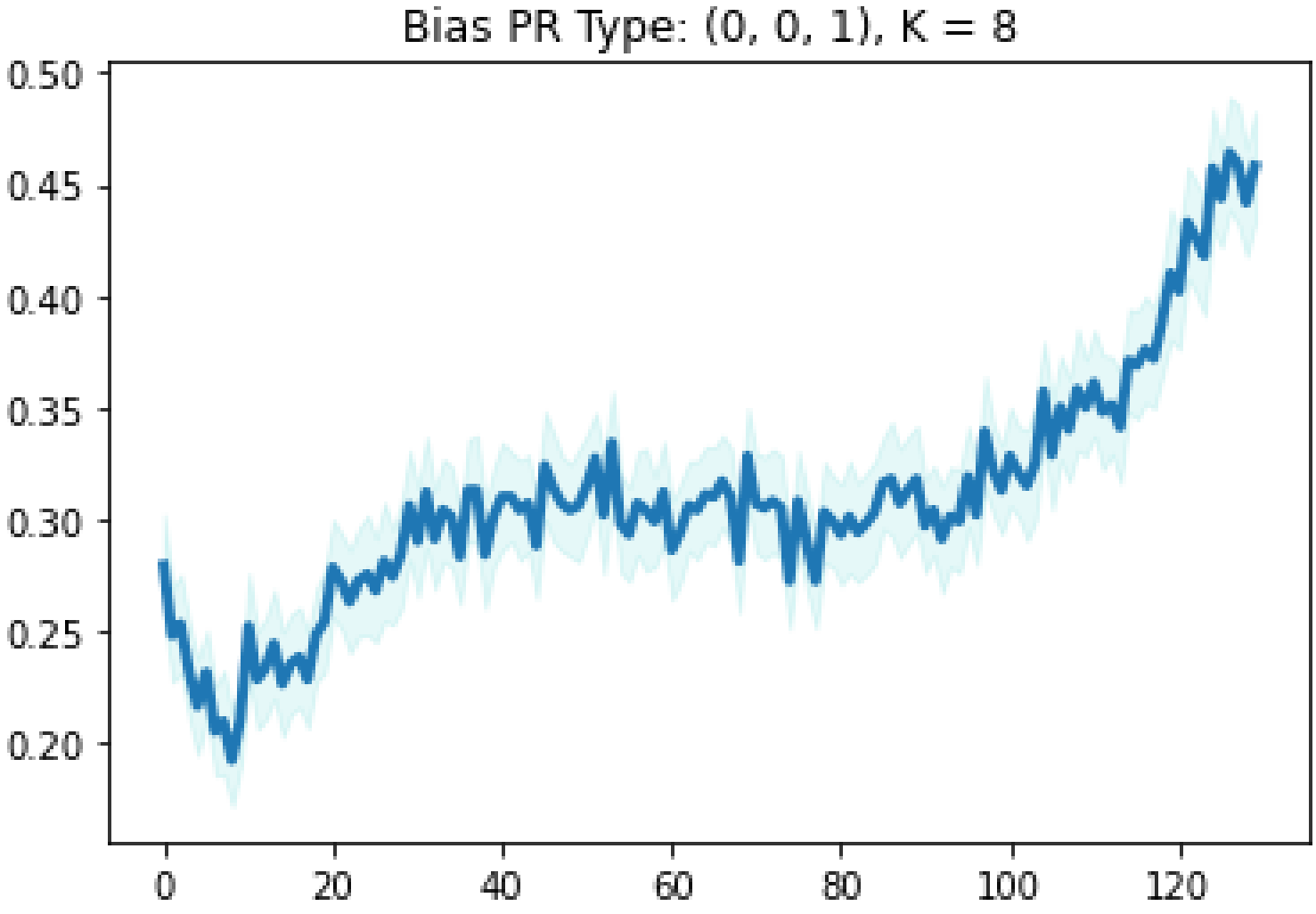
Preliminary results - DB, Reliable reviewers

Reliable = q_e drawn from a normal distrib. centered on q_m, disturbed by different epistemic approaches



Preliminary results - Status

Status = q_e drawn from a normal distrib. centered on $q_m \times o$ for those in top x% of resources, disturbed by different epistemic approaches



To Sum Up

Main insight: the lifecycle of the research area matters

Reviewers' strategies/behaviors might impact the speed of knowledge accumulation (how fast a field develops)

Limitations and future developments

Current limitations:

- Random movement
- Random reviewer matching
- No scientific collaboration (e.g. co-authorship)
- Process without memory (e.g. past publications)

Planned developments:

- Incorporate the role of the editor
- Ground the NK fitness function

Thanks for your attention

carlo.debernardi1@unimi.it,francesco.renzini@unimi.it

UNIVERSITÀ
DEGLI STUDI
DI MILANO

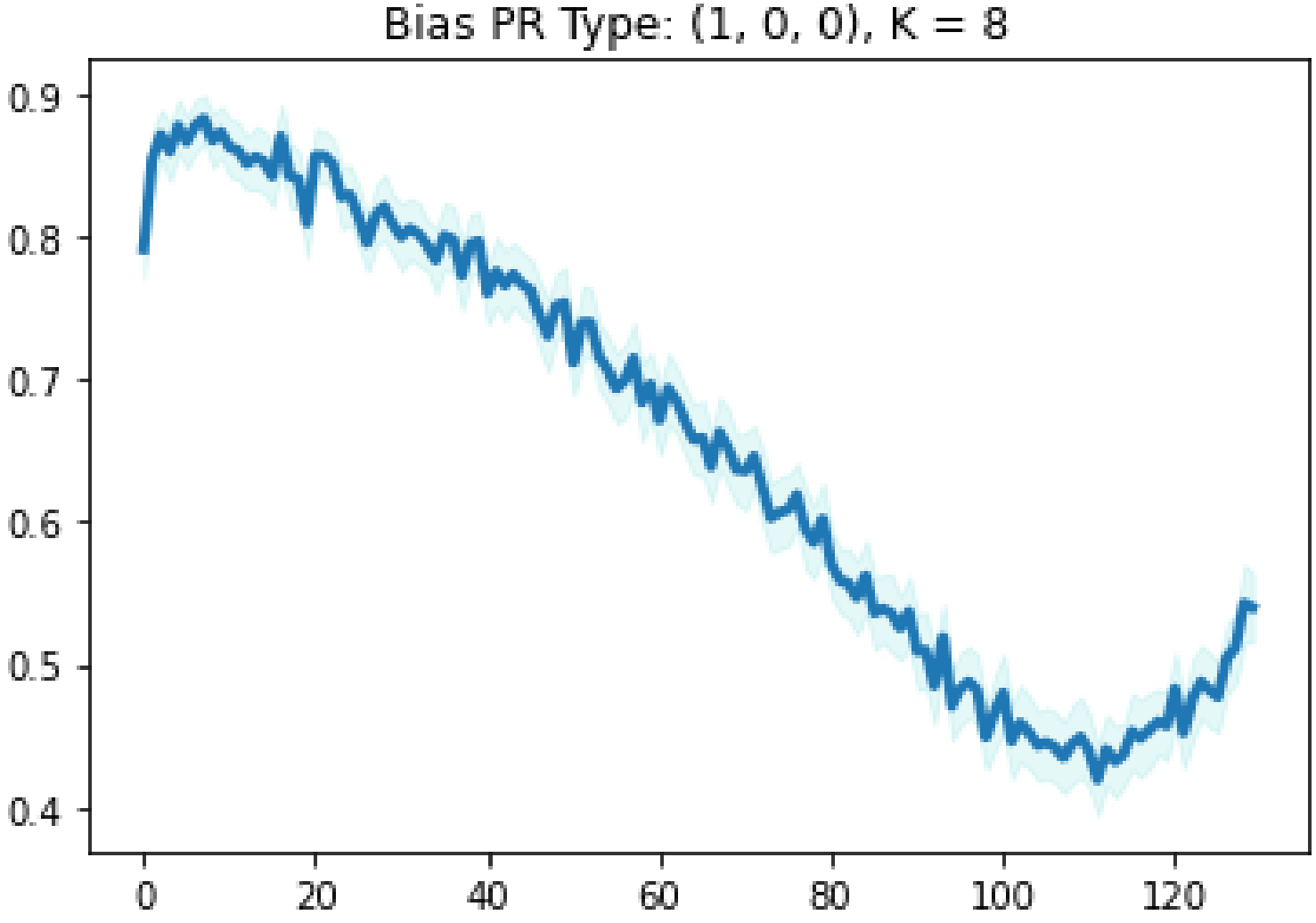
UiA Universitetet
i Agder

BEHAVE

Appendix

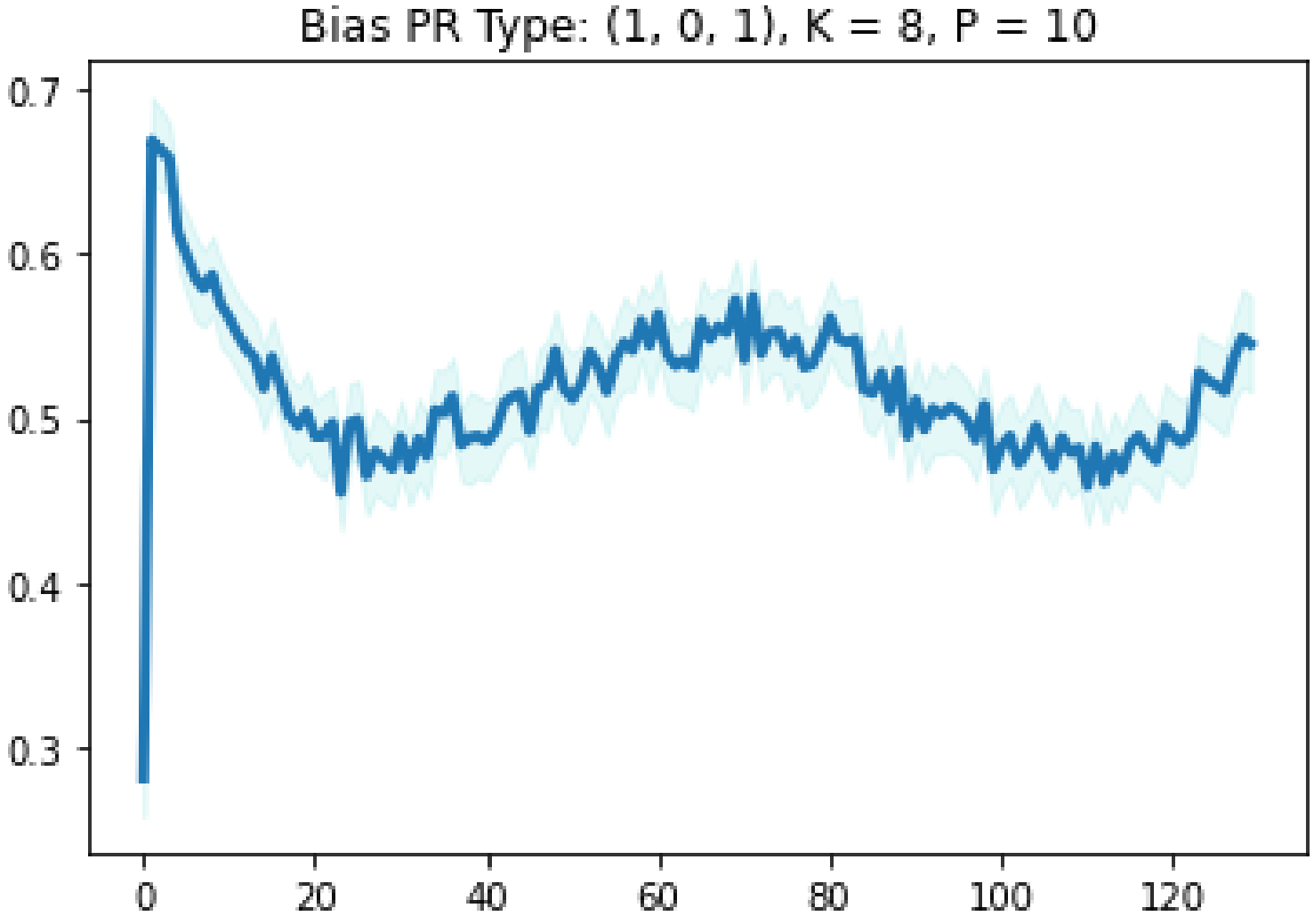
Preliminary results - Indirect Reciprocity

IR = q_e drawn from a normal distrib. centered on $q_m \times o$, disturbed by different epistemic approaches



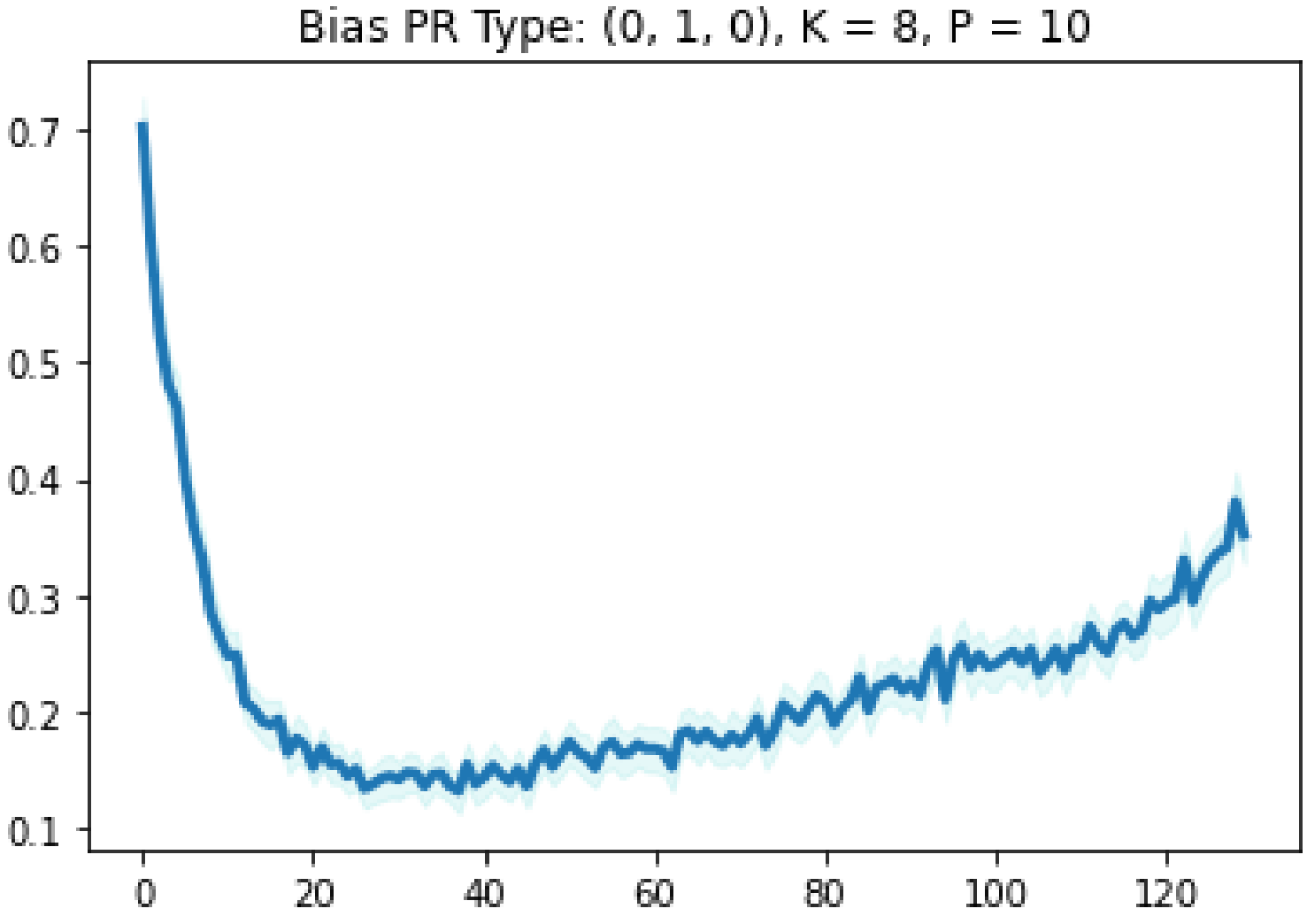
Preliminary results - Status + IR SB & Open

Status + IR = q_e drawn from a normal distrib. centered on $q_m \times o$ for those in top x% of resources + IR, disturbed by different epistemic approaches



Preliminary results - DR Open

DR $= q_e$ drawn from a normal distrib. centered on $q_m \times o$, disturbed by different epistemic approaches; thinking o too small



Preliminary results - DR Open

$o = 3.5$

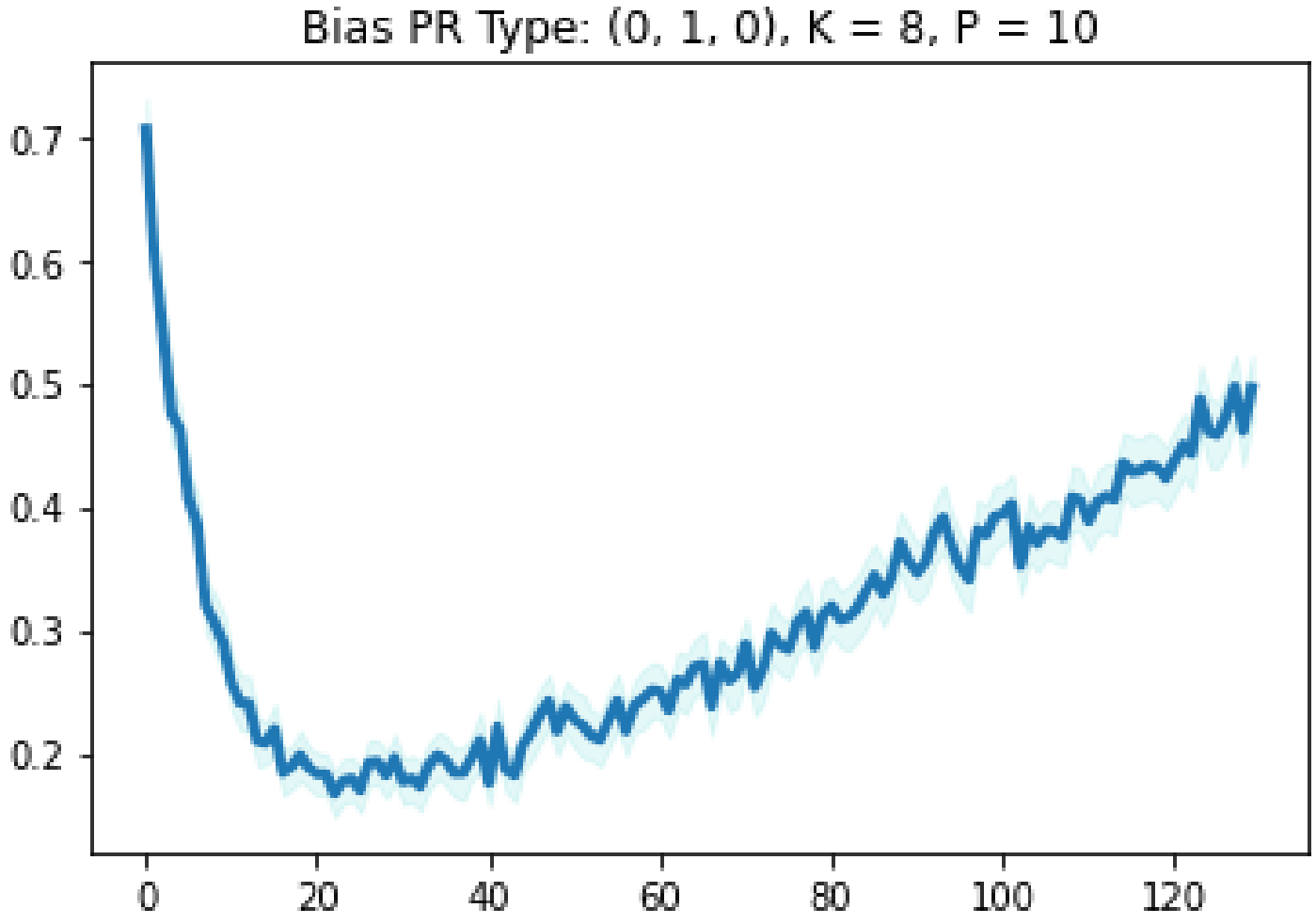


Figure: Publication bias DR Open, $o = 3.5$

Preliminary results - Status + DR + IR Open

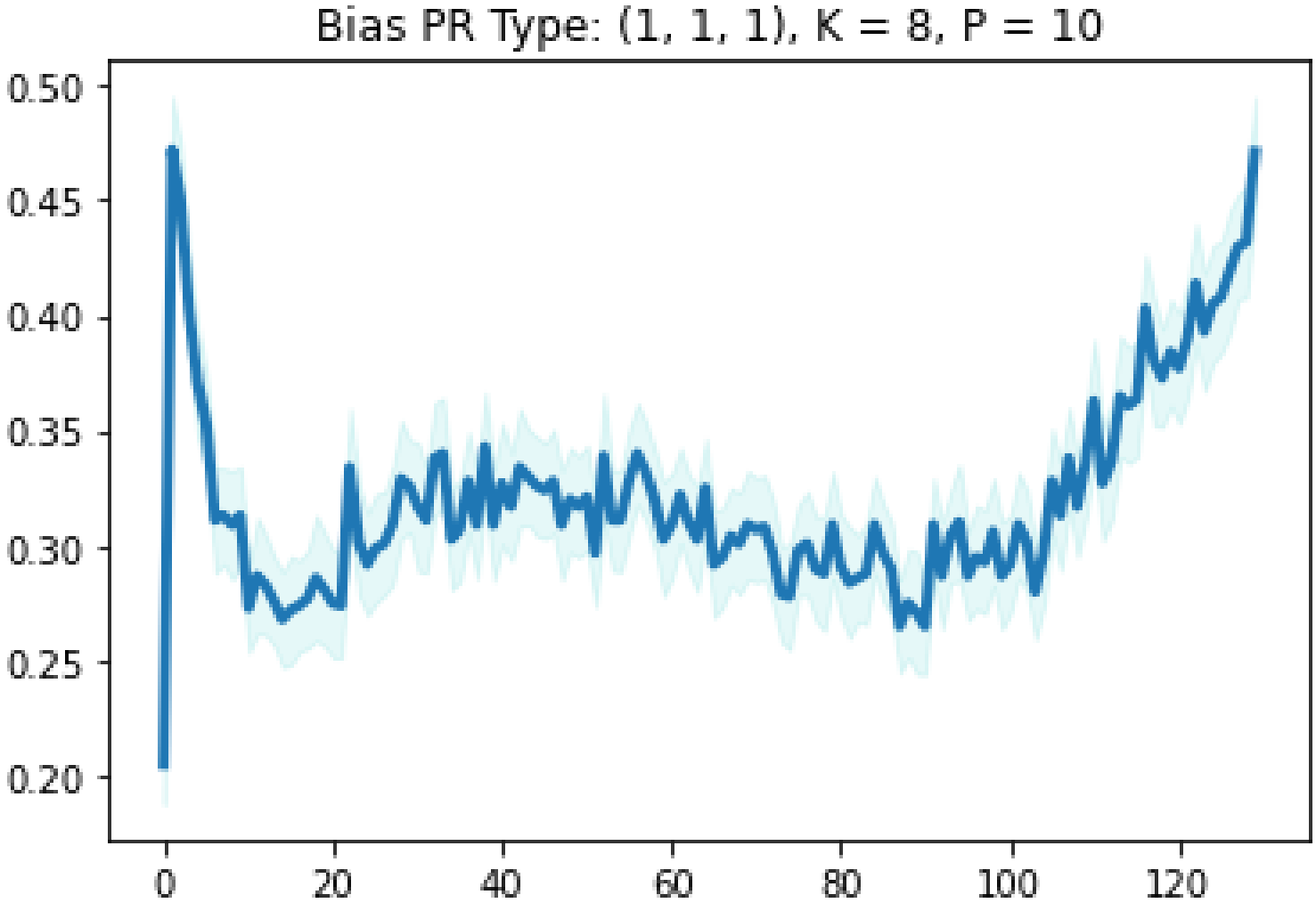


Figure: Publication bias Status + DR + IR Open