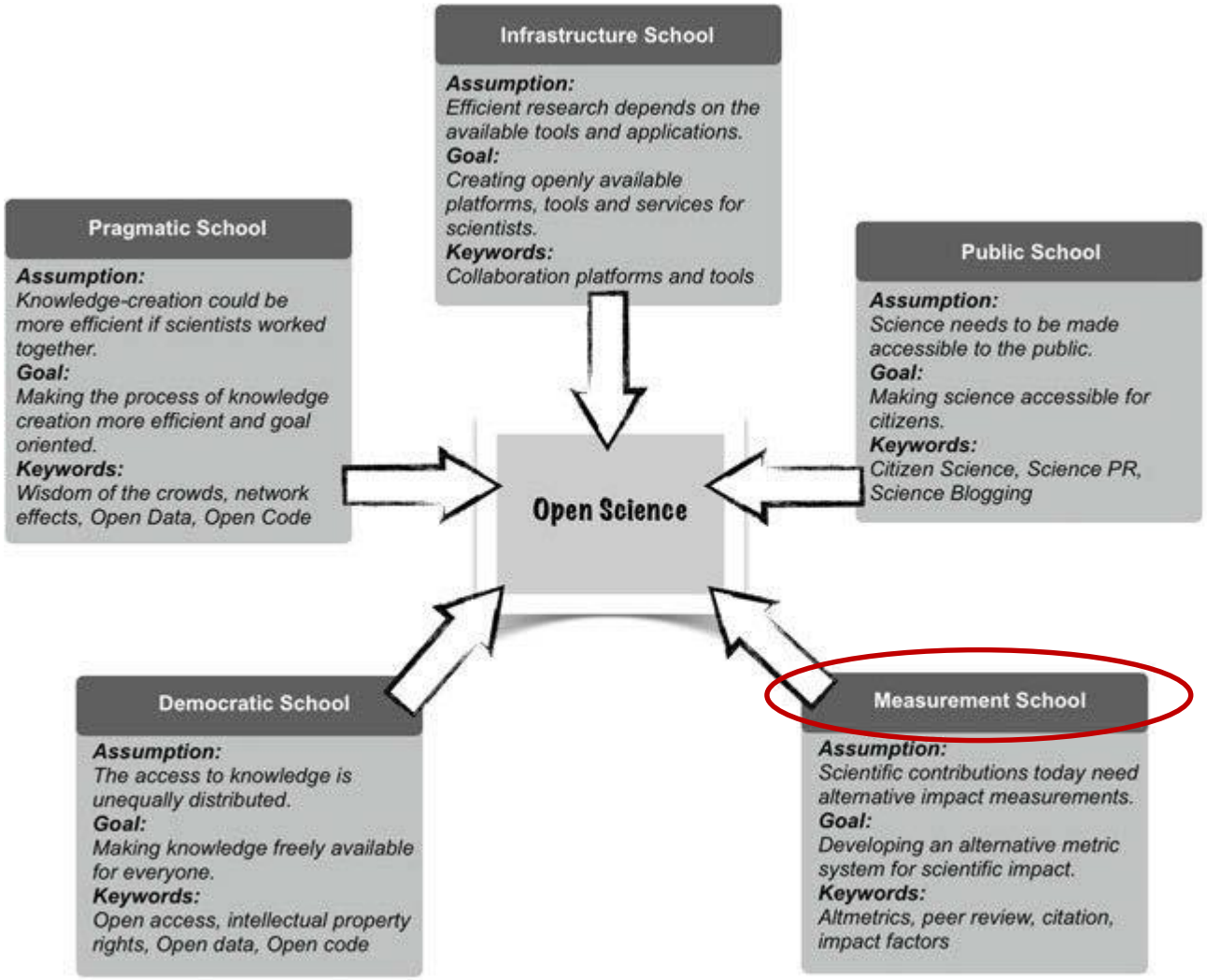
LIBRARY AND INFORMATION CENTER OF
THE HUNGARIAN ACADEMY OF SCIENCES
DEPARTMENT OF SCIENCE POLICY AND SCIENTOMETRICS

Bibliometrics as a key to Open Access Publishing – but not the way you think of it

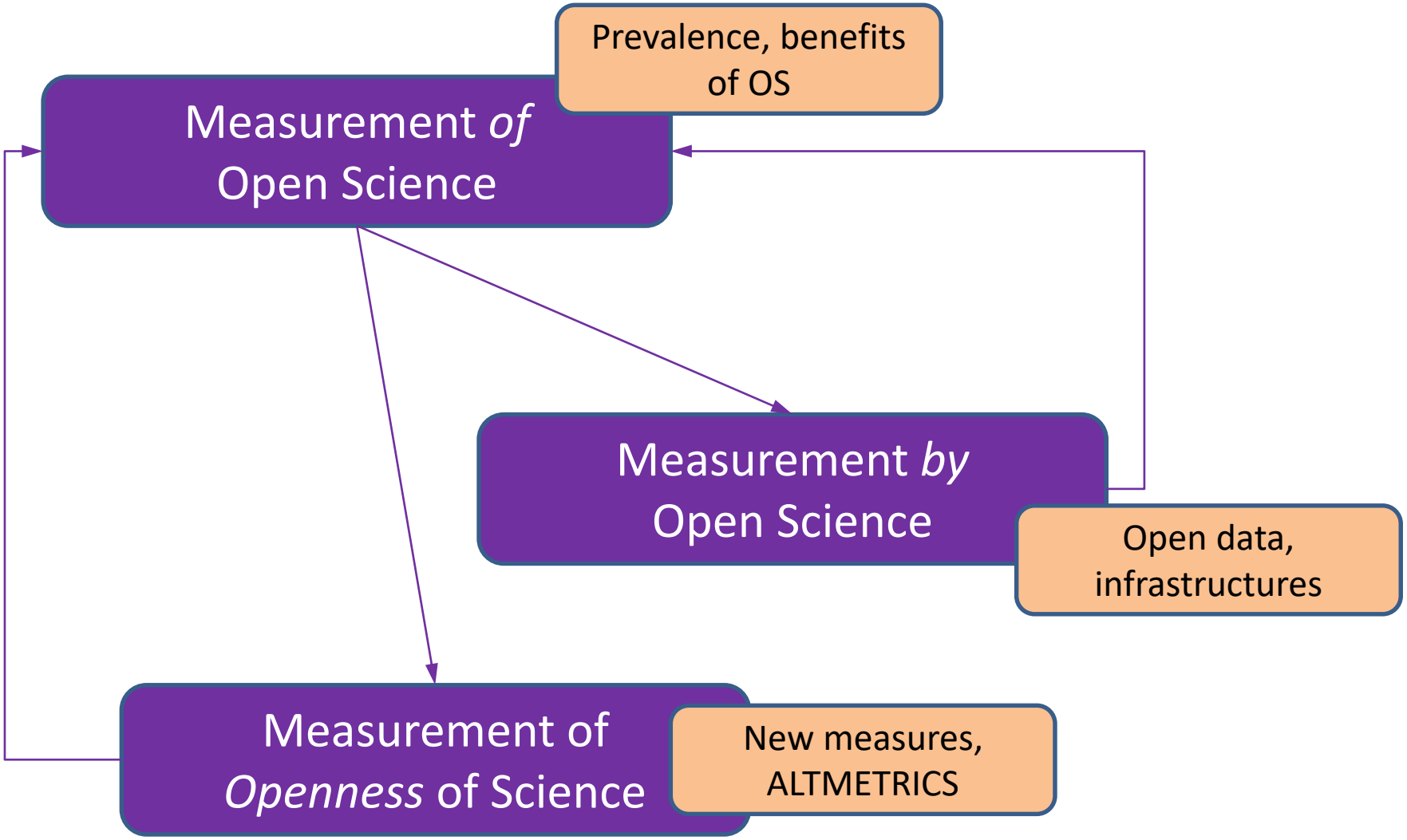
Sándor Soós

Open science schools of thought



Fecher, B., & Friesike, S. (2014). Open science: one term, five schools of thought. In Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing, Springer Open, 17-47.

Measurement and Open Science



The bibliometric measurement of OS (OA)

Open Access Citation Advantage (OACA)

"T*Although they have been cited in support of OA mandates by institutions and funders, industrial applications [1], charity [2] and public appeal [3] are not in themselves sufficient to motivate the adoption of such mandates, or compliance with them.* ***Research impact itself would seem to be the natural rationale for maximizing research access****."*

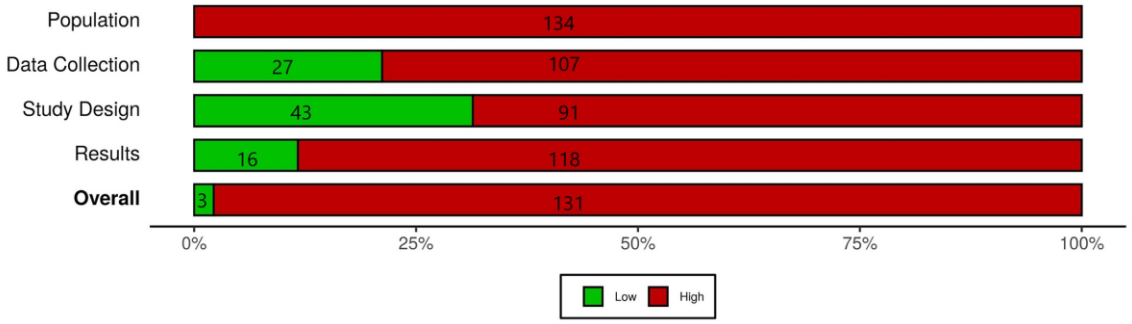
Harnad, S. (2008). Confirmation bias and the open access advantage: some methodological suggestions for the Davis citation study. *arXiv preprint arXiv:0808.3296*.

Open Access Citation Advantage (OACA)

Research synthesis: Systematic reviews

Study Design		
	Randomized	2 (1.5%)
	Non-Randomized	132 (98.5%)

Does Open Access Citation Advantage Exist?			
Yes (n = 64)	No (n = 37)	Sometimes (n = 32)	Inconclusive (n = 1)



Langham-Putrow, A., Bakker, C., & Riegelman, A. (2021). Is the open access citation advantage real? A systematic review of the citation of open access and subscription-based articles. *PloS one*, *16*(6), e0253129.

Tendencies:

- OACA typically found in observational, not found in experimental studies (causation!)
- Critique of experimental: short-term, small and context-dependent samples
- Even observational is mostly context-dependent

The “MDPI conundrum”

International trend:

- The industrialization of OA publishing → Market of scholarly communication
- Business model → “anomalies” of publishing (journals, publishers – accused of – questionable, predatory (?) practices)
- Common attitude: results in lower scientific value and quality often disguised (below standard peer review, inflated metrics) → hinders quantative performance assessment (= bibliometric evaluation)
- Demonised: MDPI and its mega-journals (typically *Sustainability*)

"If two journals have the same impact factor, but one of them is part of Springer Nature and the other one is part of MDPI, it is a rather different quality.“

(Personal opinion communicated at a CoARA WG meeting).

The “MDPI conundrum”

Adverse effects on the academic publishing culture:

- Proliferation of criticism on dominant forms of OA → Revival old and initial OA scepticism (lower quality, substandard etc.)
- Case study from Hungary: *MTA’s recommendations on new types of publication misconduct**
  - Triggers for this declaration: (1) concerns on the research assessment climate forcing the acceleration of Q1-level publishing in relation to (2) the “inflated metrics” opinion, (2) circumstacial evidence: steep increase in MDPI share withon country output
  - General stance (simplified): avoid publishing in MDPI (Frontiers, Plos)
  - Adverse effect: renewal of the atmosphere of OA-suspicion (APC-based, but overgeneralization is evident) [anecdotal evidence]
- Most fundamental problem:

  New types of (1) publication misconduct vs. (2) publication culture?

- *https://mta.hu/english/proposals-for-the-handling-of-articles-for-journals-that-engage-in-objectionable-practices-mtas-recommendations-on-new-types-of-publication-misconduct-113312

The “MDPI conundrum”

Can bibliometrics (measurement) do justice?

- Option 1: New (bibliometric) measures for scholarly performance
- Option 2: Bibliometric analysis of the behaviour and characteristics of the “questionable” venues and research published through them

Hint: Option 2 is more viable

- Bibliometric – empirical – analysis can, among many other things
  - (1) evaluate anecdotal evidence on publication venues (e.g. on “OA is suspicious”, “IF of MDPI journals is less a proxy of journal prestige than editorial policy”, “MDPI papers are less cited or only self-cited at publisher/journal level” etc and (2) represent the drivers of publishing
  - (2) Reveal the factors of publishing behaviour (e.g. choice of venues) and the impacts of such interventions as the MTA recommendation/declaration
  - Go beyond descriptive statistics on OA publishing

- In what follows, two such analyses are briefly demonstrated

The “MDPI conundrum”

The “quality” conveyed by conventional journal metrics (for journals)

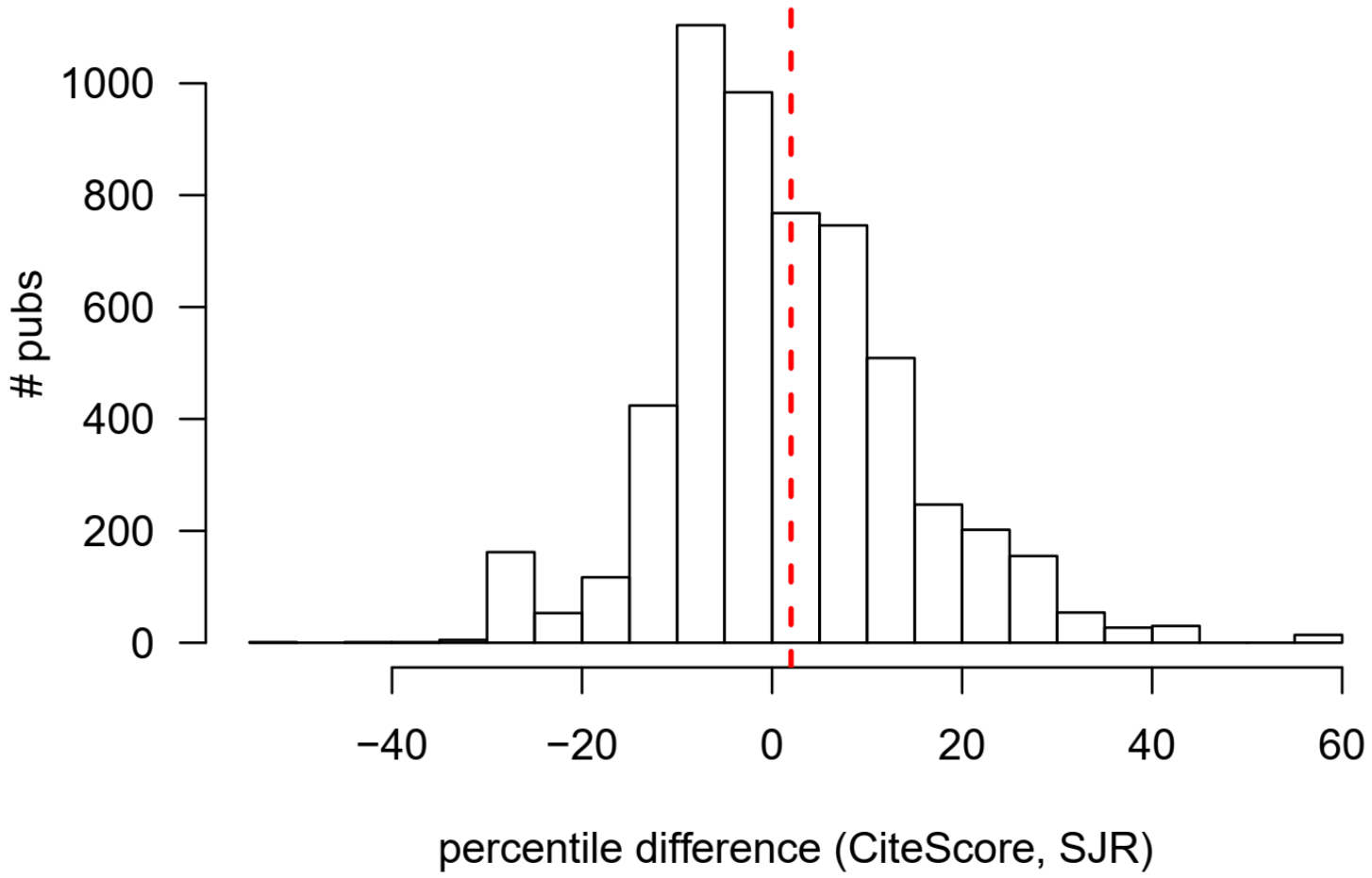
- **RQ**. Is the citation count-based (JIF-like) journal metrics overrepresent the prestige of MDPI journals (inflating JIF-like metrics with citations from less recognized, more peripheric journals)?

- **Method.** Comparative analysis of count-oriented and prestige-oriented metrics for a large-scale output (CiteScore vs. SJR, rank-based comparison).

- **Data.** Country-level publication output in Scopus, 2018—2022 (5-year pub window)

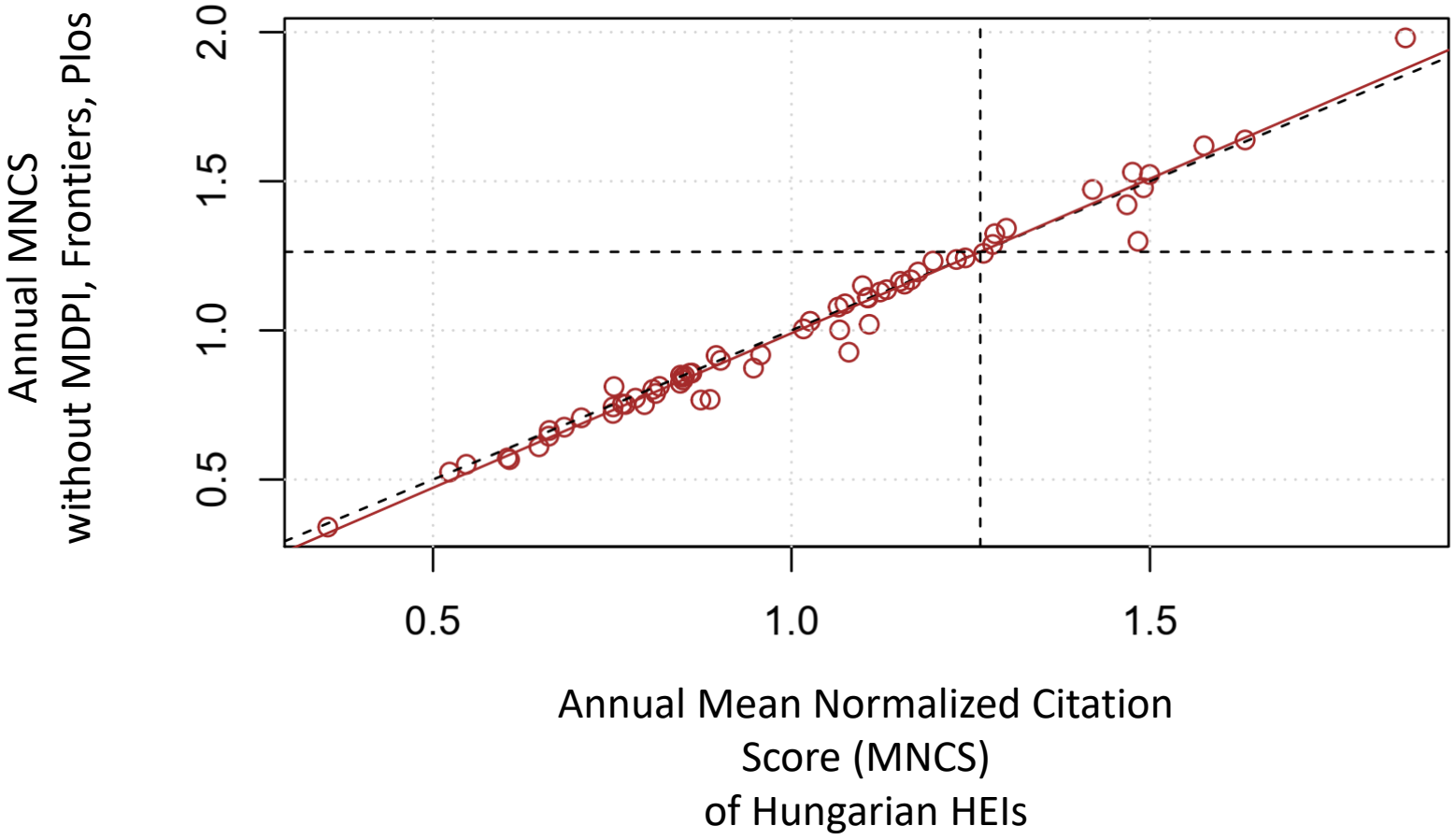
The “MDPI conundrum”

The “quality” conveyed by conventional journal metrics (for journals)



The “MDPI conundrum”

Context dependence: evidence from Hungary



The “MDPI conundrum”

Identification of an explanatory model of publishing in MDPI

- **RQ1**. What factors influence the choice of publication venues, specifically: gold OA vs. other, more specifically: MDPI vs. other? What is the role (contribution) of factors?
- **RQ1.1.** Is there an effect attributable to an intervention?
- **Related works.** Surveys on researcher attitudes towards OA publishing (but lack of bibliometric approaches)

- **Method.** Study of theoretically relevant bibliometric variables as predictors of publication patterns for (*collaborating*) authors. **Intervention is conceived as the MTA recommendations** effective from mid-2023 – effect is modelled as the effect of publishing years on venue choice.

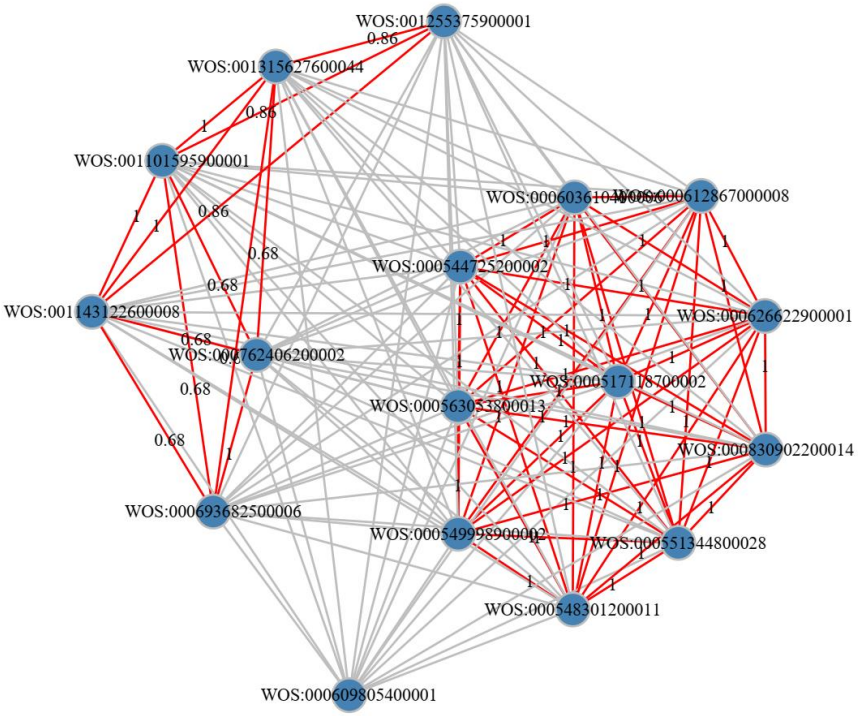
- **Data.** Country-level publication output in WoS, 2020—2024 (~5-year pub window). (~60K pubs, 200K authors)

The “MDPI conundrum”

Identification of an explanatory model of publishing in MDPI

- **Novelty**. A predictor is constructed to account for the effect of author clusters (accustomed to venues)
- **Variable:** *Author clusters.* Technically: clusters of papers with similar author profiles

- **Theoretically sound model:**
- Unit of analysis: paper. The model is to predict its publisher type (MDPI or else) based on predictors as
- *Author cluster*, Author characteristics (*international*, *domestic*, *number of authors*), *Year*, *Journal metrics*, controls (*Reaearc Area* etc.)



The “MDPI conundrum”

Identification of an explanatory model of publishing in MDPI

- **Simplified model (for testing the effect of intervention)**
- The test of annual differences in MDPI-share within the individual paper conglomerates conveying author groups (*Author clusters* and *Year* as predictor)

Paired Samples T-Test

			statistic	df	p	Mean difference	SE difference		Effect Size
2024	2023	Student's t	-1.36	318	0.176	-0.0433	0.0319	Cohen's d	-0.0760
2023	2022	Student's t	1.21	318	0.227	0.0402	0.0332	Cohen's d	0.0678
2022	2021	Student's t	2.19	318	0.029	0.0625	0.0286	Cohen's d	0.1225
2021	2020	Student's t	1.64	318	0.102	0.0418	0.0255	Cohen's d	0.0917

Note. H_a $\mu_{\text{Measure 1 - Measure 2}} \neq 0$

Summary

How to “unlock the potential of (biblio)metrics?”

- It is not necessarily *alternative constructs*, *alternative measures* and *alternative data* what is needed.
- New *applications* of bibliometrics: totally existent but scarce in this context: structural (not evaluative) measurement
- Benefit: exploring and evidencing the inclusion of OS within scholarly communication

Thank you for your attention!