

SciVal and Scopus User Meeting, UK & Ireland, October 2022

Evidencing Open Access Citation

Advantage using SciVal

Evidencing Open Access Citation Advantage

2

OA Citation Advantage:

- *the relative citation success of our open work vs. our non-open work*

Motivation:

- replace anecdote with evidence, for use in OA advocacy

And along the way:

- gather insight into our authors' publication practice
- anticipate the return on investment of our JISC Transitional Agreements

Method *

1. acquire an entity's list of publications via the Overview module
2. split this list by OA type
3. create Publication Sets for each type
4. use the Benchmark module to compare OA types

We mainly use My SciVal groups to define the entities.

Unhelpful features in SciVal

Overall research performance



All Open Access?

- Rather: *free to access (at the point of observation)*
- Bronze is very useful for insight: but it isn't OA re: prevailing OA policy

Green OA is non-exclusive

- includes items found in other OA types
- Not incorrect, but not helpful for our analysis

Alternative routes to publication sets

5

Route 1

- Export entity's publication list
- Drop DOIs into [Unpaywall Simple Query Tool](#)
- Use Excel to filter and select by 'oa_status'
- Drop DOIs for each type into the import wizard for publication sets

OR

Route 2

- Use filters on the entity publication list to create publication sets for Gold, Hybrid gold, Bronze and Not OA
- Export 'Open Access' attribute & EIDs for the Green OA filter
- Use Excel 'text to columns' to get exclusively green items
- Drop EIDs for green items into the import wizard for publication sets

Why use unpaywall?

Unpaywall simple query tool result:

doi	best_oa_evidence	best_oa_host	best_oa_licen	best_oa_url	best_oa_url_i	best_oa_version	oa_status
10.1093/mnras	open (via free articl	publisher		https://acade	TRUE	publishedVersion	bronze
10.1088/1361-4	open (via crossref li	publisher	cc-by	https://doi.o	TRUE	publishedVersion	hybrid
10.1093/mnras	oa repository (via O	repository		http://uhra.h	TRUE	publishedVersion	green
10.1016/j.scitotenv.2018.08.410					FALSE		closed
10.1051/0004-4	open (via crossref li	publisher	cc-by	https://www	TRUE	publishedVersion	hybrid
10.1186/s1226	oa journal (via doaj	publisher	cc-by	https://genes	TRUE	publishedVersion	gold

VS

SciVal export result:

10.1016/j.comppsy	Gold Green	2-s2.0-85084572248
10.1051/0004-6361	Bronze Green	2-s2.0-85012880358
10.1109/COMST.20	Green	2-s2.0-85042127896
10.1051/0004-6361	Bronze Green	2-s2.0-85060375915
10.1016/j.jacc.2016	Hybrid gold Green	2-s2.0-85012974187

- The additional data available in unpaywall evidences the algorithm that calculates 'os_status'
- License info is useful to extend the analysis (27% cc-by; 15% cc-by-nc; the remainder largely unspecified)

Unpaywall vs SciVal data compared

7

	Scival	Unpaywall	Diff' (S-U)
Entity size	3950	3783	
no doi (SciVal)	167		4% ¹
no result (Unpaywall) ²		42	?
Gold	773	803	-30
Hybrid Gold	441	449	-8
Green	1448	1459	-11
Bronze	486	443	43
Closed ²	593	587	6
Matched	3741		
Agree	3661	98%	
Disagree	80	2%	
OA	2662	2711	
Not OA	1079	1030	
OA rate (matched)	71%	72%	-1%
OA rate (SciVal entity)	67%	69%	-2%

- ¹ SciVal has 4% of items (with only 0.5% citations) that do not have a DOI (majority conference contributions)
- ² Unpaywall oa_status is specific about 'closed' and 'no result'; whereas SciVal just leaves the Open Access attribute empty when there is no open outcome.
- There is 98% agreement over 3.8K items
- The 2% disagreement can be characterised by saying that SciVal reports slightly less Gold, Hybrid gold and Green; and slightly more bronze and no status.
- The 80 disagreeing items have 1.3% of citations, and lead to 1% underestimate in SciVal in the headline OA rate.
- Conclude: the differences between the two systems are insignificant for large entities

Benchmarking #1 – set the scene

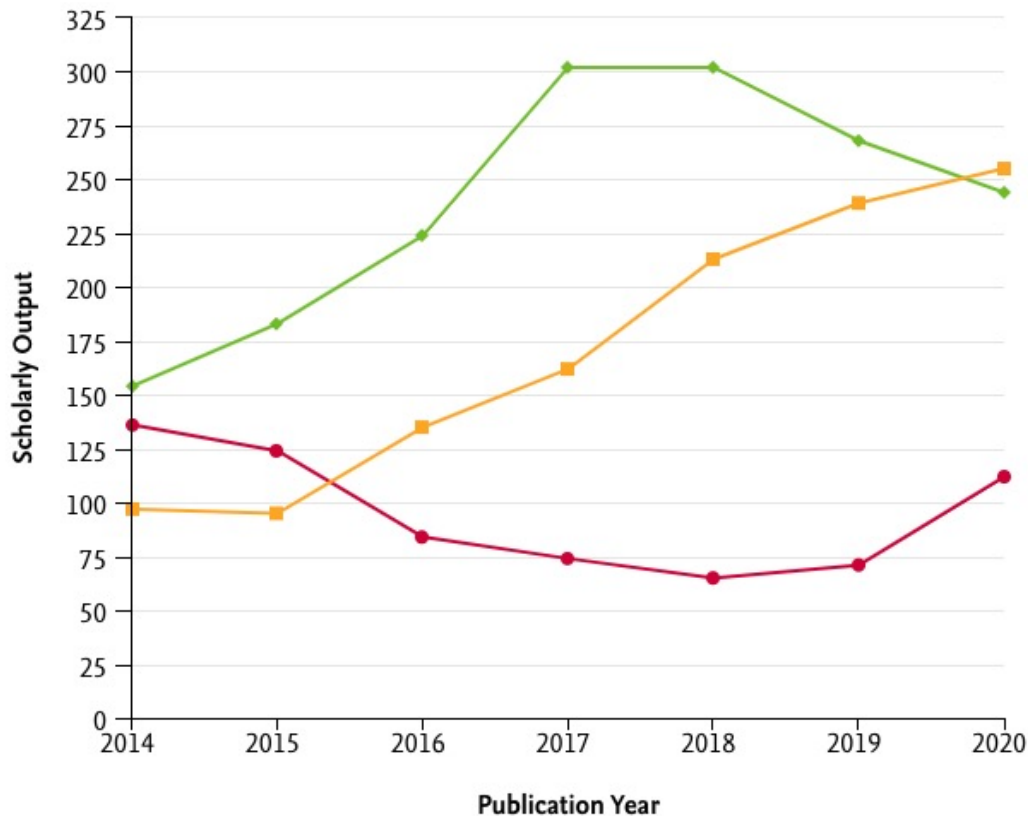
y-axis

x-axis

Bubble size

Scholarly Output

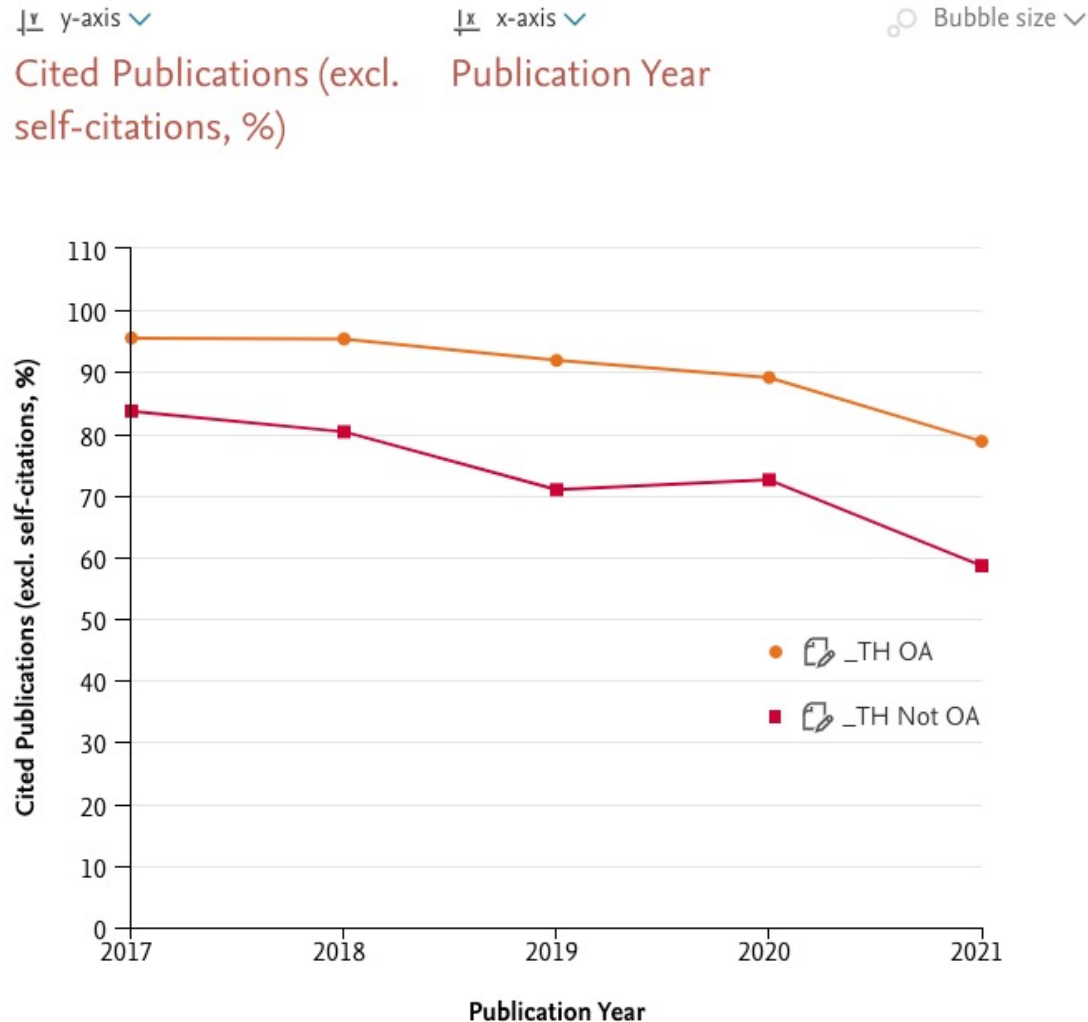
Publication Year



- Green OA 2014-2020 (Articles + Reviews)
- Gold OA 2014-2020 (Articles + Reviews)
- Not OA 2014-2020 (Articles + Reviews)

- Rise of Green OA in response to REF2021
- Gold OA gradually catches up

Benchmarking #2 – blunt instrument?



- Whatever your story remind the audience : metrics are indicators, not facts; and; we make no reflection on the quality of the work
- Rather than saying “*there is an open access citation advantage*”, perhaps it is better to say “*the ‘Not OA’ work may have been ‘disadvantaged’ by its inaccessibility*”?
- Or just state the evidence: *for this set of publications* citation advantage = 90% cited for OA vs. 73% cited for not OA

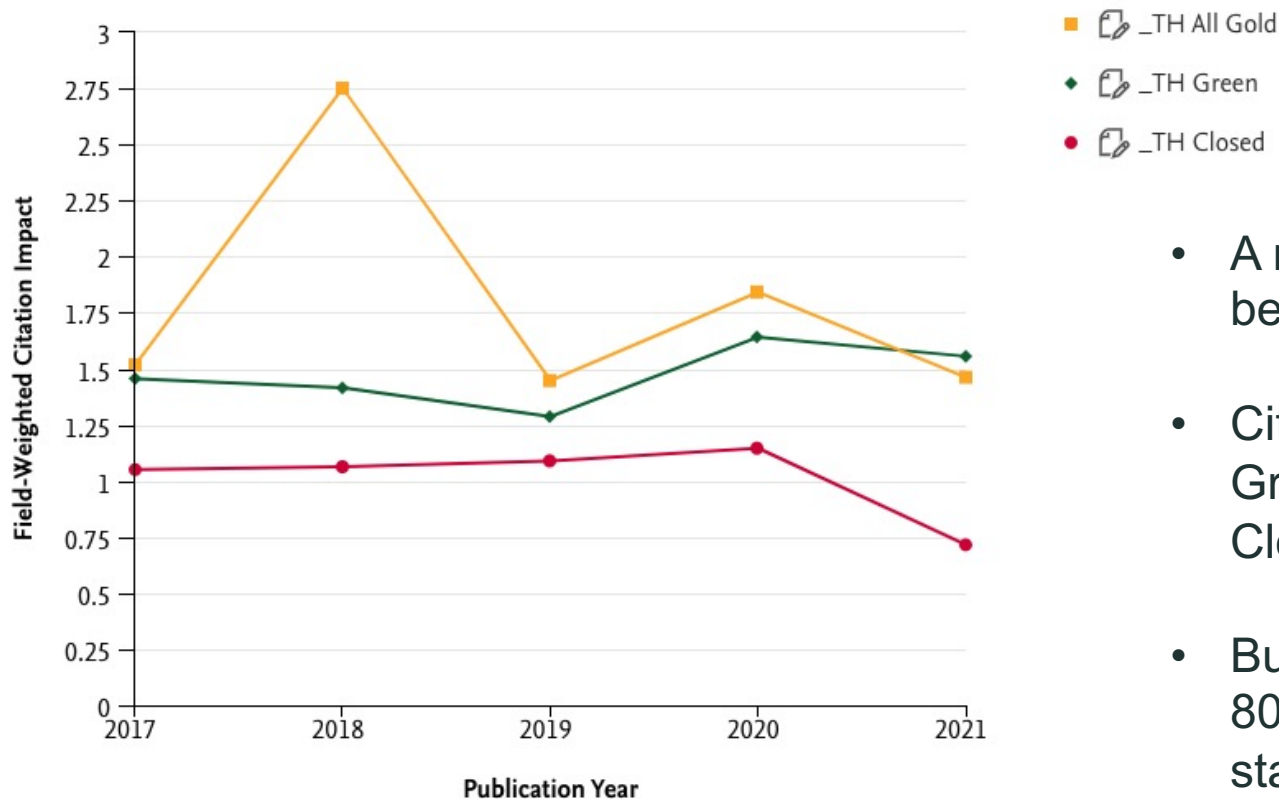
Benchmarking #3 – a better OA conversation?

10

y-axis x-axis Bubble size

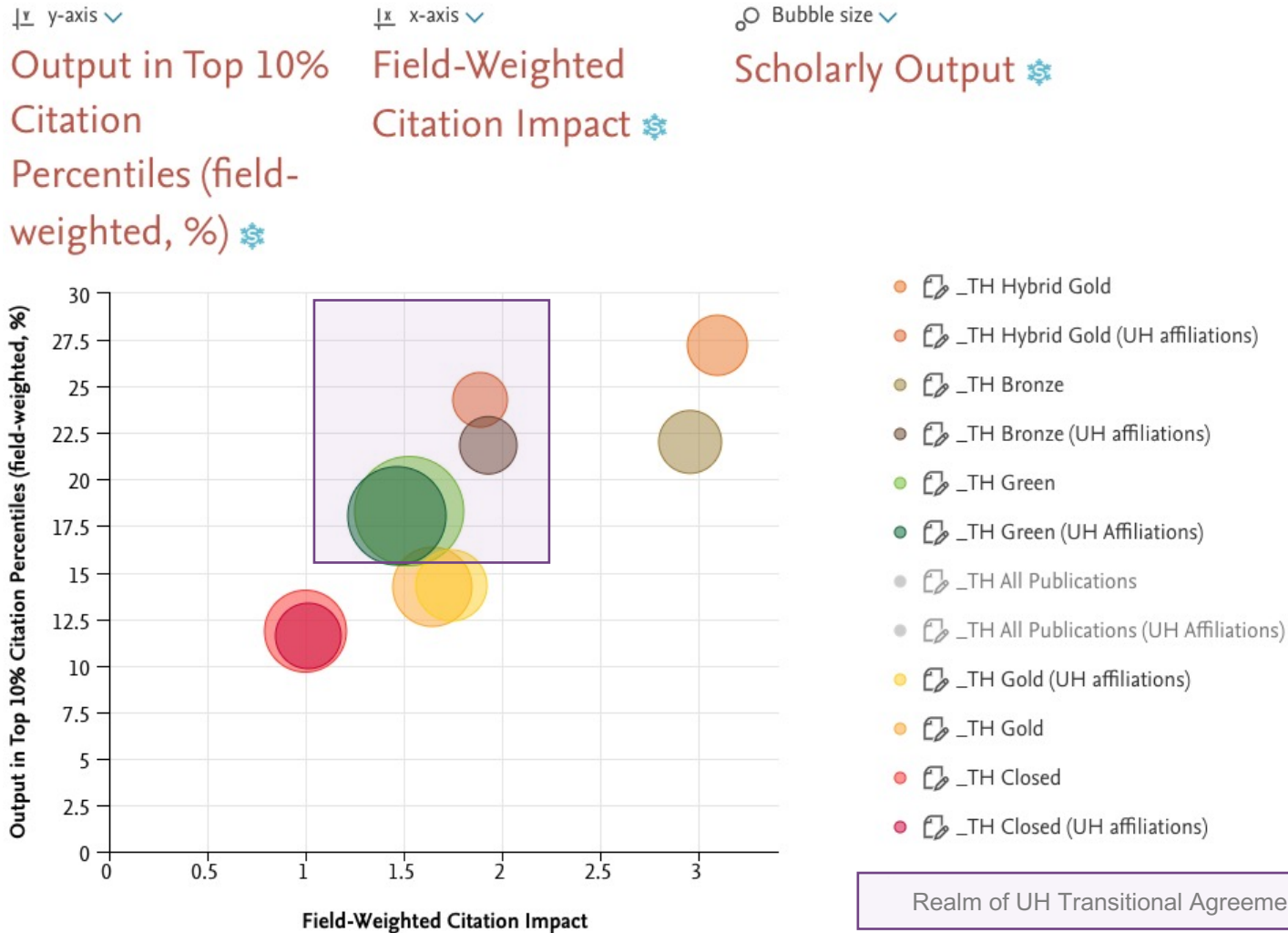
Field-Weighted Citation Impact Publication Year

Impact



- A more detailed and positive message with better hooks into the OA conversation?
- Citation advantage: FWCI 0.5 higher for Green OA, and 0.8 higher for Gold OA vs Closed.
- But note the 2018 spike. Just paper 1 (out of 800) - a hyper-authored standard practice statement with FWCI ~300 - has significantly distorted the metric.

Benchmarking #4 – packing it all in?



- almost a basket of snowball metrics in one visualisation!
- Citation advantage: ~0.7 greater FWCI, and nearly double outputs in top percentiles for OA vs Closed.
- ‘Closed’ obviously disadvantaged
- Note the performance of Bronze
- Note Hybrid advantage over native Gold
- Watch this space for a JISC Transitional Agreement ‘dividend’ as Green and Bronze convert to Gold?

Thank you; Q & A