2018 STM IGNITE SESSION

WTF? Predictive Analytics
Better Data – Better Tools

Dan Pollock, Director of Data and Analytics
Dan.Pollock@deltathink.com
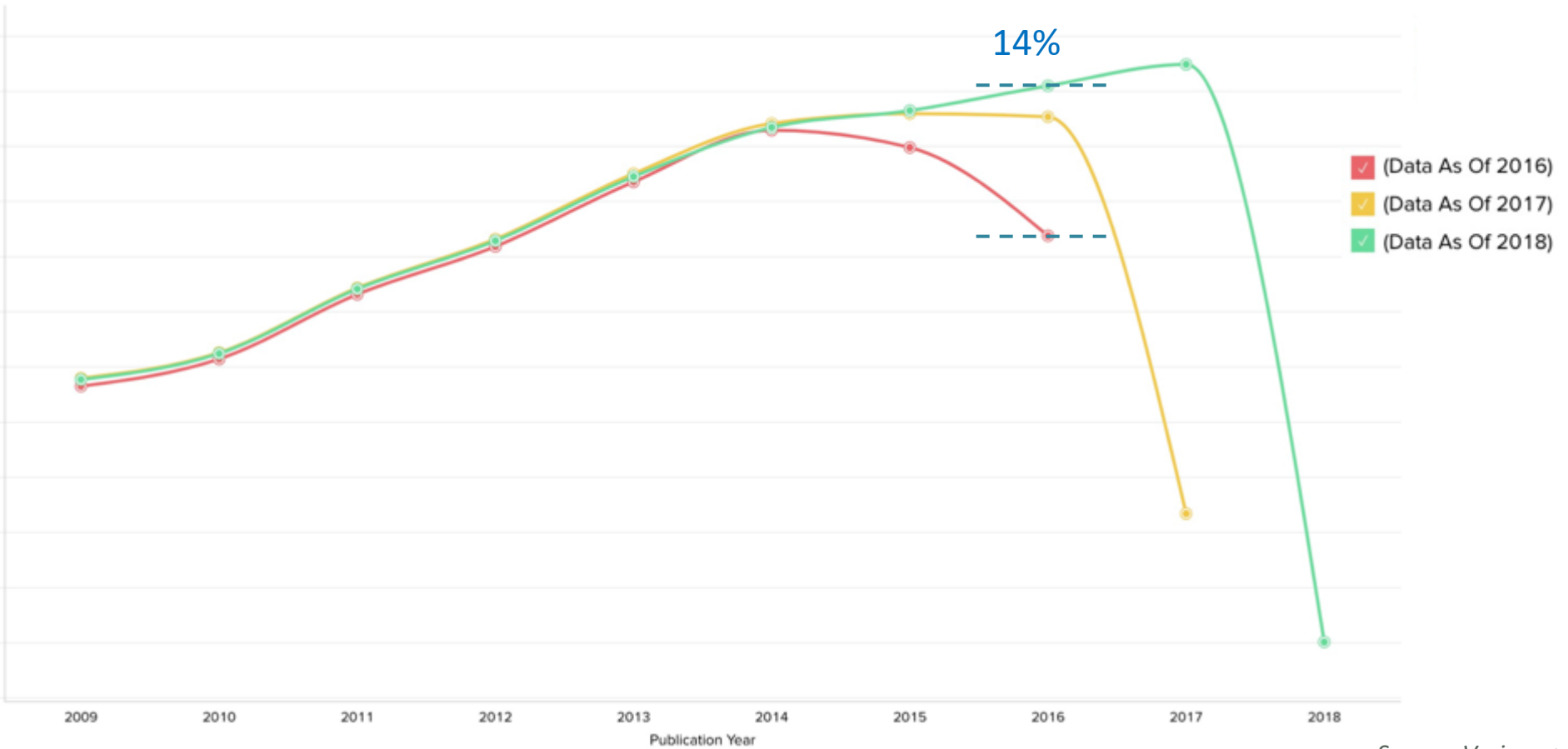
Delta Think

Data: A Neutral View

Data Tools WTF?

Data Tools

WTF?

What's The Future

U.S. Recorded Music Revenues by Format

1973 to 2017, Format(s): All
Source: RIAA

Legend:
- LP/EP
- Vinyl Single
- Cassette
- Cassette Single
- 8 - Track
- Other Tapes
- CD
- CD Single
- SACD
- DVD Audio
- Music Video (Physical)
- Download Album
- Download Single
- Download Music Video
- Ringtones & Ringbacks
- Other Digital
- Kiosk
- On-Demand Streaming (Ad-Suppo...
- Other Ad-Supported Streaming
- SoundExchange Distributions
- Limited Tier Paid Subscription
- Paid Subscription
- Synchronization

Total # Scholarly Articles Published – Data Variance

14%

☑ (Data As Of 2016)
☑ (Data As Of 2017)
☑ (Data As Of 2018)

Publication Year

Source: Various +
Delta Think analysis.

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$$

Where β_0, β_1 and β_2 are parameters to be estimated from the data. Standard practice is to find values of these parameters such that the sum of squares:

$$\sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_i + \beta_2 x_i^2) \right]^2$$

is minimized. In words, we are looking for coefficients of the polynomial such that the fitted values of the polynomial are as close to the observations as possible. In matrix/vector notation what we want is the vector $\vec{\beta}$ which satisfies:

$$\vec{y} = X\vec{\beta}$$

where $\vec{\beta} = [\beta_0, \beta_1, \beta_2]^T$, $\vec{y} = [y_1, \cdots, y_n]^T$ and

$$X = \begin{bmatrix} 1, x_1, x_1^2 \\ 1, x_2, x_2^2 \\ \cdots \\ 1, x_n, x_n^2 \end{bmatrix}$$

As we cannot invert the matrix X (it's not square for one thing), we solve the equation as follows:

$$X^T \vec{y} = X^T X \vec{\beta}$$
$$(X^T X)^{-1} X^T \vec{y} = \vec{\beta}$$

Q

What are the implications of Plan S?
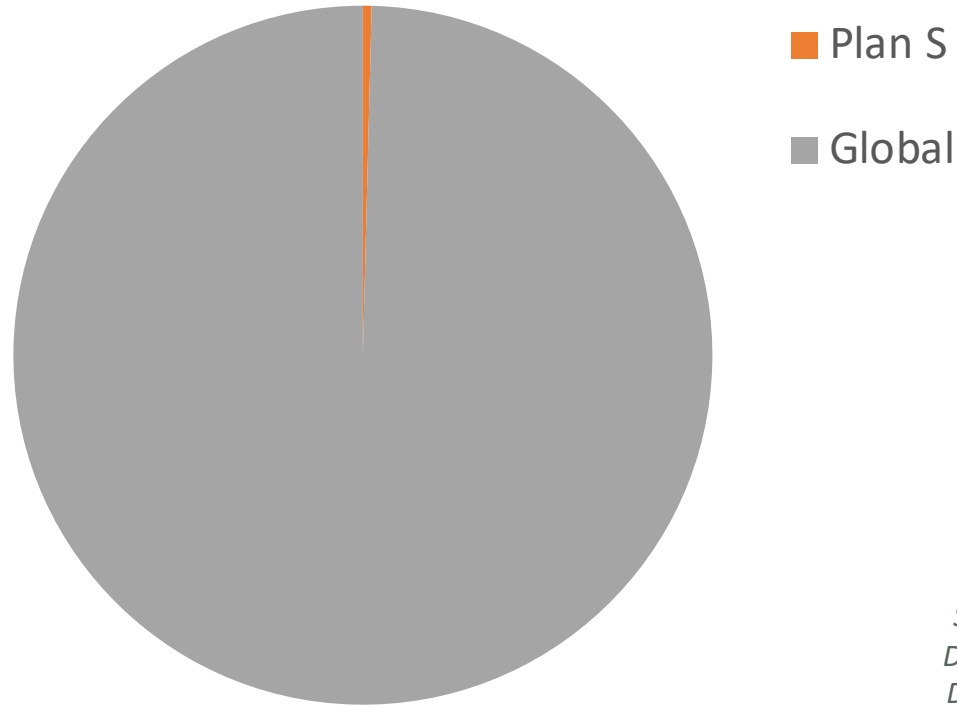
Plan S Funders' Share of Global Expenditure on R&D

- Plan S
- Global

Source: R&D Magazine, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.

http://deltathink.com/ sign-up-for-news-views/

Plan S Funders' Share of Global Article Output
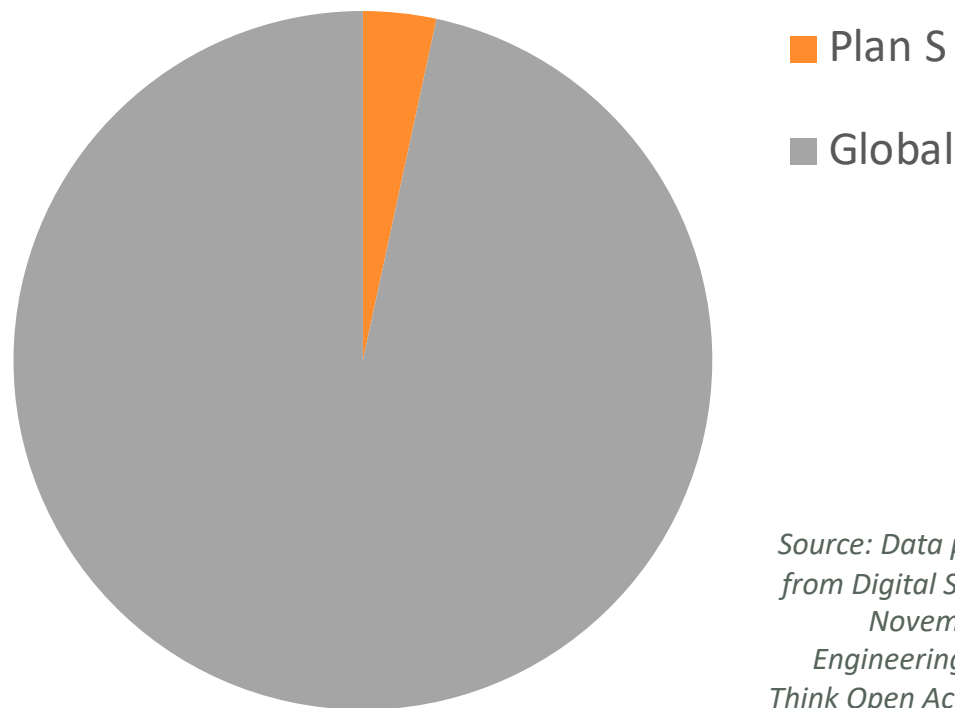
■ Plan S

■ Global

Source: Data provided by Dimensions from Digital Science - collected on 16 November 2018, US Science & Engineering Indicators 2016, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.

http://deltathink.com/ sign-up-for-news-views/

Plan S: share of scholarly articles in context

Scholarly articles in 2017...	Shares of Global Research Articles
Plan S Funders Share of Global Output	3.5%

http://deltathink.com/ sign-up-for-news-views/

Plan S: share of scholarly articles in context

Scholarly articles in 2017...	Shares of Global Research Articles	Share of Global OA Output
Plan S Funders Share of Global Output	3.5%	4.5%

**http://deltathink.com/
sign-up-for-news-views/**

Plan S: share of scholarly articles in context

| Scholarly articles in 2017...	Shares of Global Research Articles	Share of Global OA Output
Plan S Funders Share of Global Output	3.5%	4.5%
Plan S Funders' Share of Global Output including equivalent level of funder coverage from Germany	4.2%	5.3%

Source: Data provided by Dimensions from Digital Science - collected on 16 November 2018, US Science & Engineering Indicators 2016, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.

http://deltathink.com/ sign-up-for-news-views/

Change in Market **Value** of Plan S Uptake Scenarios Compared with Current Projections

Change in Market **Value** of Plan S Uptake Scenarios
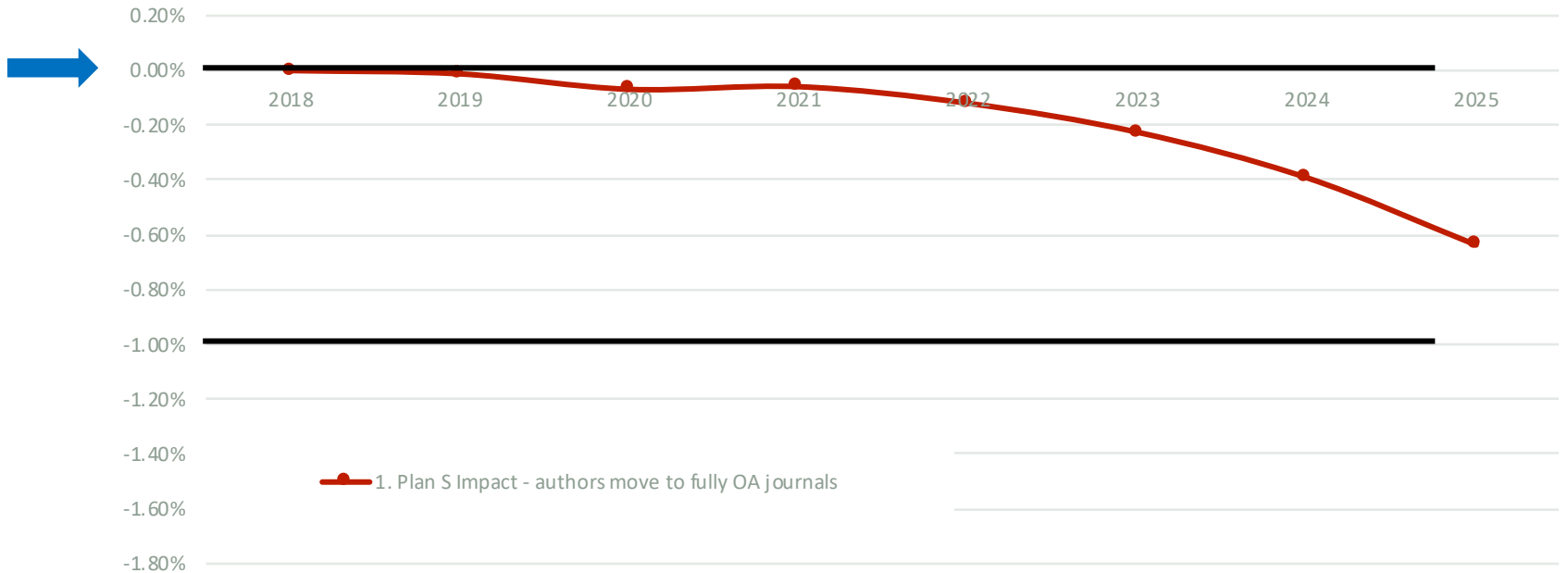
Source: Delta Think Open Access Data and Analytics Tool, Dimensions from Digital Science, Delta Think analysis.
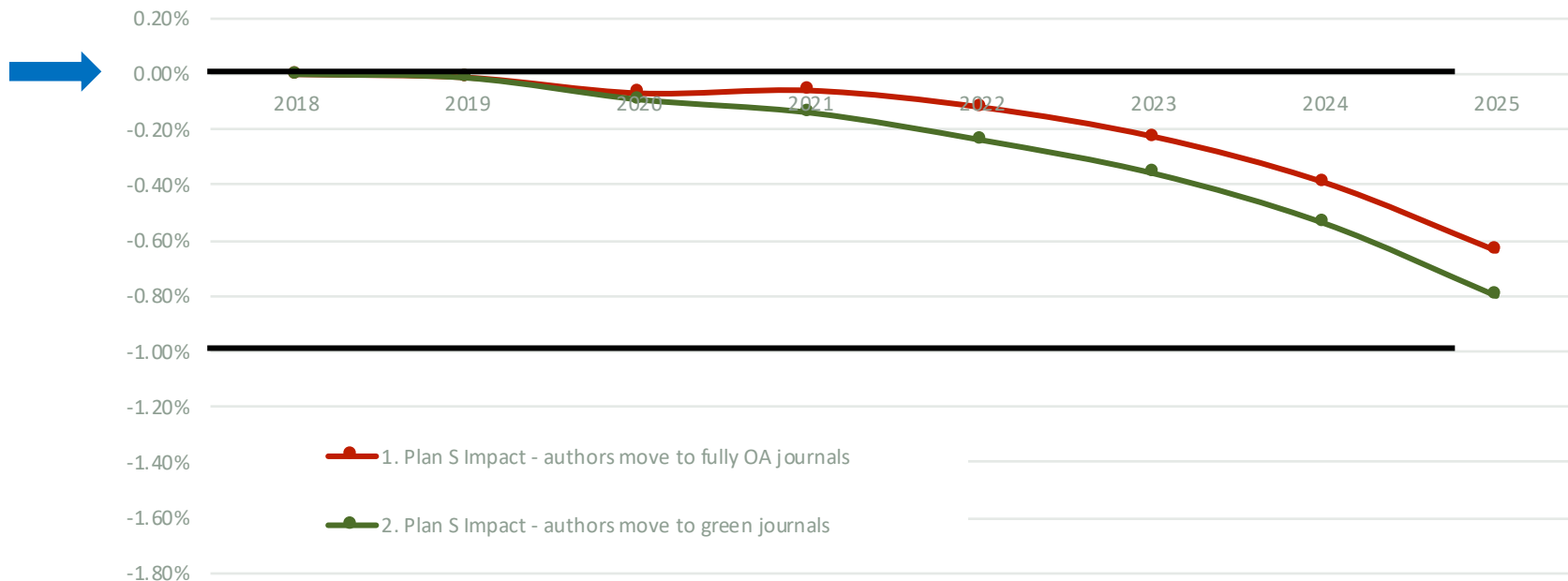
Change in Market **Value** of Plan S Uptake Scenarios
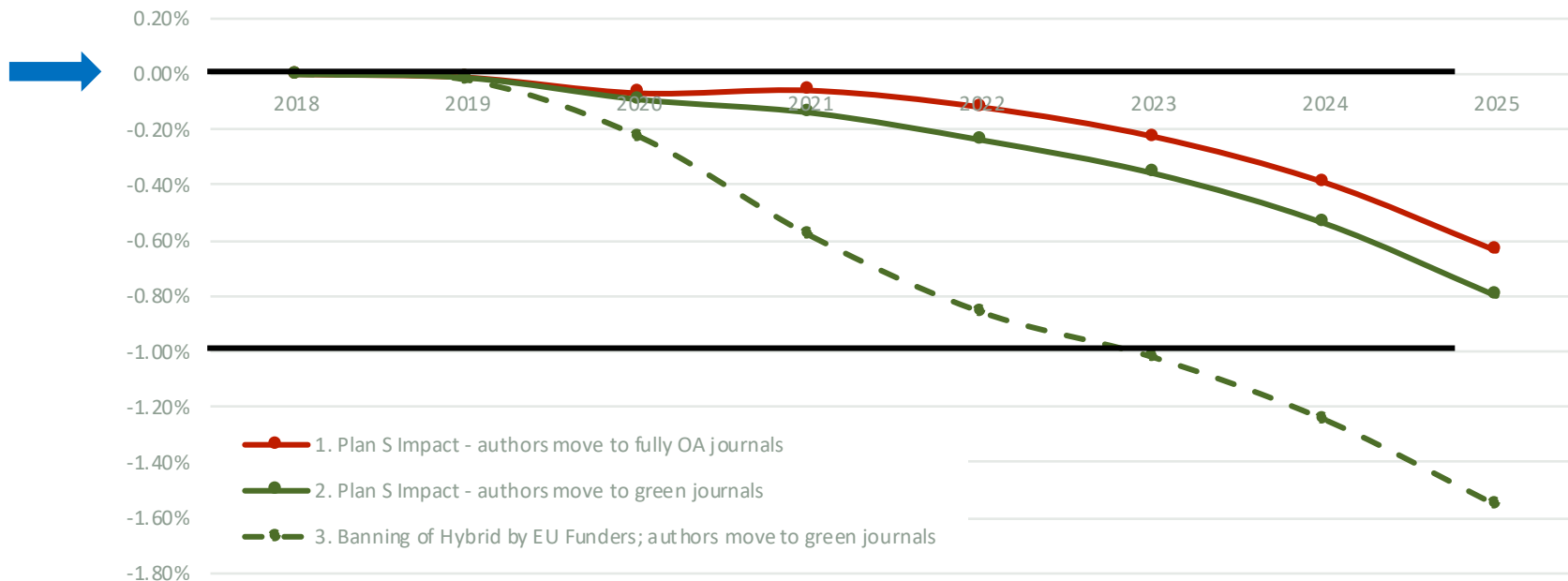

Source: Delta Think Open Access Data and Analytics Tool, Dimensions from Digital Science, Delta Think analysis.

Change in Market **Value** of Plan S Uptake Scenarios Compared with Current Projections

Legend:
- 1. Plan S Impact - authors move to fully OA journals
- 2. Plan S Impact - authors move to green journals
- 3. Banning of Hybrid by EU Funders; authors move to green journals

Q

Does this data stuff really matter?

https://www.elsevier.com/about/open-science/open-access/surprising-facts

Fact 2:

Our APC prices are value for money

Elsevier's article publishing charges (APCs) are comparable with other publishers.

We publish the best quality open access articles (based on field-weighted citation impact) for both open access and hybrid titles, while our APCs remain at or below industry average.

Our APC price list is publicly available and downloadable here.

At Elsevier we do not charge subscribers for open access publications and only take into account subscription content when pricing subscription titles.
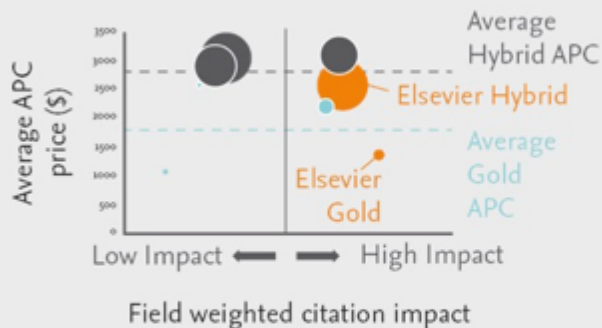
Key

🔵 Gold Competitor
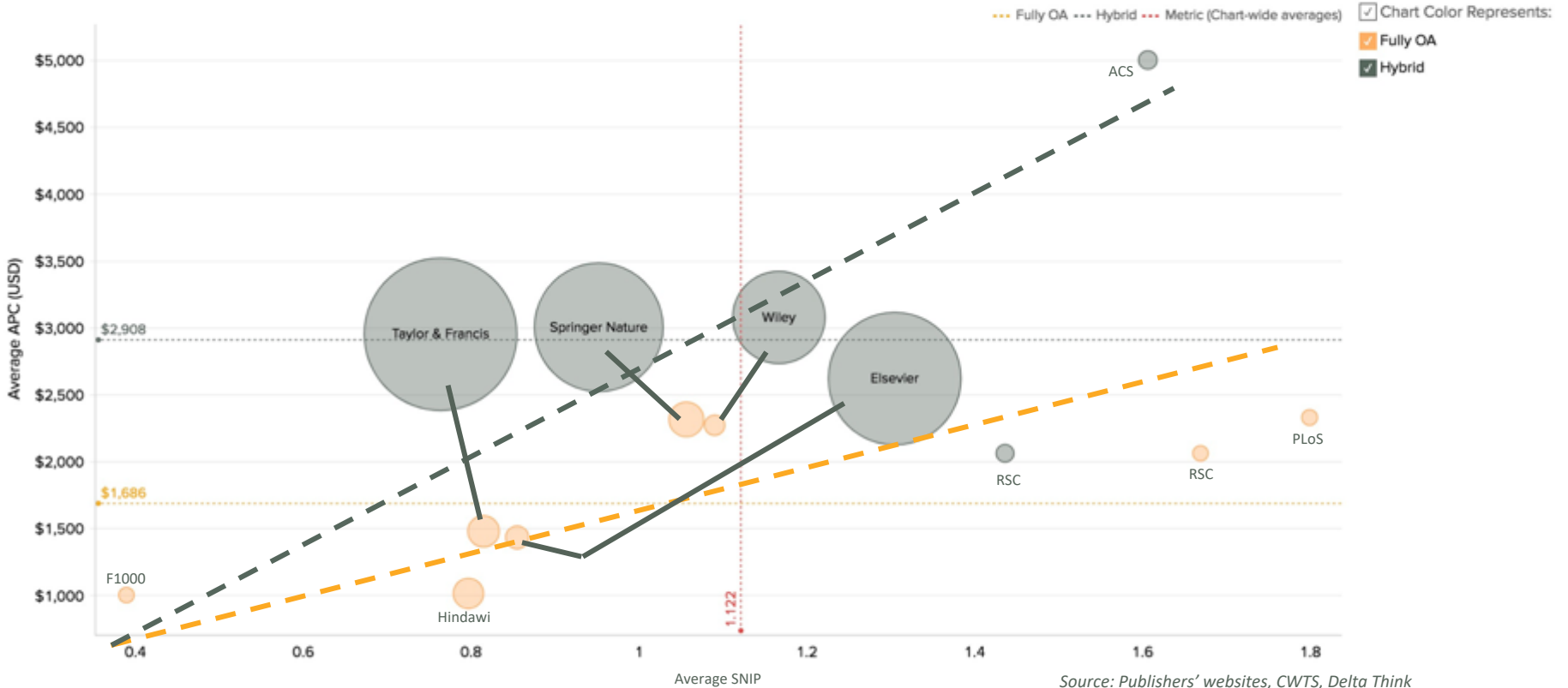⚫ Hybrid Competitor

ESAC – Efficiency and Standards for Article Charges

- Collating information about agreements
- Common approaches to negotiation
- Data driven

…an example of changes in approach?

Market position
Share of global article output 2016, open access gold share

Portfolio overview
Open access gold and hybrid shares by the number of articles

Proprietary & Confidential

Pricing vs. Impact Metrics per Publisher ⟳

METRIC:
SNIP ▾

Color By:
Fully OA/Hybrid ▾

Pricing Year:
2017-18 ▾

Zero or Blank APC::
Non-zero APC ▾

Fully OA / Hybrid:
All ▾

Publisher:
Elsevier,Taylor & ▾   Reset

···· Fully OA ··· Hybrid ··· Metric (Chart-wide averages)

☑ Chart Color Represents:
☑ Fully OA
☑ Hybrid

Average APC (USD)

$5,000
$4,500
$4,000
$3,500
$3,000 — $2,908
$2,500
$2,000
$1,686
$1,500
$1,000

ACS
Taylor & Francis
Springer Nature
Wiley
Elsevier
RSC
PLoS
RSC
F1000
Hindawi

1,122

0.4   0.6   0.8   1   1.2   1.4   1.6   1.8

Average SNIP

Challenges

- Common OA data needs

- Nebulous OA data sources

- Neutral data sets

- Complex calculations

OA DAT: A One-Stop Shop

- Address the challenges

- *Accessible answers*

Dan.Pollock   Emma.Green
@deltathink.com   @deltathink.com

**http://DeltaThink.com/
sign-up-for-news-views/**

Delta Think