OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Determinants and prediction of APC price levels

Workshop for OA data experts | OpenAPC Workshop
Bielefeld, 17.01.2019

Dr. Nina Schönfelder

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Agenda

- Analysing the determinants of APC-levels

  Schönfelder, N. (2018). APCs — Mirroring the impact factor or legacy of the subscription-based model? Bielefeld: Universitätsbibliothek Bielefeld. doi:10.4119/unibi/2931061

- Projecting APCs for currently hybrid or closed-access journals

- Comparing projected total APC-spending with libraries budgets‘ after a hypothetical full journal flipping

  - on a global level
  - for each German university/research institute

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Data

- OpenAPC data set (from 2018-02-06)
  - APCs actually paid (in contract to catalogue prices)
  - country, period, journal type (hybrid/oa), journal title, publisher
- CWTS Journal Indicators 2016 (calculated by Leiden University's Centre for Science and Technology Studies based on the Scopus bibliographic database produced by Elsevier)
  - "source normalized impact per paper" (SNIP)
  - subject area of the journal
- Data base of the "Kompetenzzentrum Bibliometrie"
  - Web-of-Science Indices: SCI, SSCI, ISTP, ISSHP, AHCI
  - Publication output of six research-performing institutions on article-level

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Summary statistics

country

```
GBR     :24572
DEU     :14054
AUT     : 4244
SWE     : 1532
NOR     : 1171
CAN     :  929
(Other): 1240
```

institution

```
UCL                           : 4526
FWF - Austrian Science Fund: 4205
Wellcome Trust                : 3782
MPG                           : 3465
University of Cambridge       : 2044
University of Oxford          : 1506
(Other)                       :28214
```

period

```
2016    :16210
2015    :12892
2014    :11178
2013    : 3253
2012    : 1472
2017    :  905
(Other): 1832
```

publisher

```
Elsevier BV                          : 6838
Springer Nature                      : 6484
Public Library of Science (PLoS)     : 5690
Wiley-Blackwell                      : 4265
Springer Science + Business Media: 3627
Frontiers Media SA                   : 2718
(Other)                              :18120
```

journal_full_title

```
PLOS ONE                  : 4789
Scientific Reports        : 1388
New Journal of Physics    :  983
Frontiers in Psychology:  680
Nature Communications     :  630
BMJ Open                  :  437
(Other)                   :38835
```

is_hybrid

```
Mode :logical
FALSE:26755
TRUE :20987
```

SNIP

```
Min.    : 0.000
1st Qu.: 1.050
Median : 1.230
Mean    : 1.435
3rd Qu.: 1.620
Max.    :15.870
NA's    :5013
```

Subject.area

```
Health Sciences              :10616
Life Sciences                :20312
Physical Sciences            : 9462
Social Sciences & Humanities: 2339
NA's                         : 5013
```

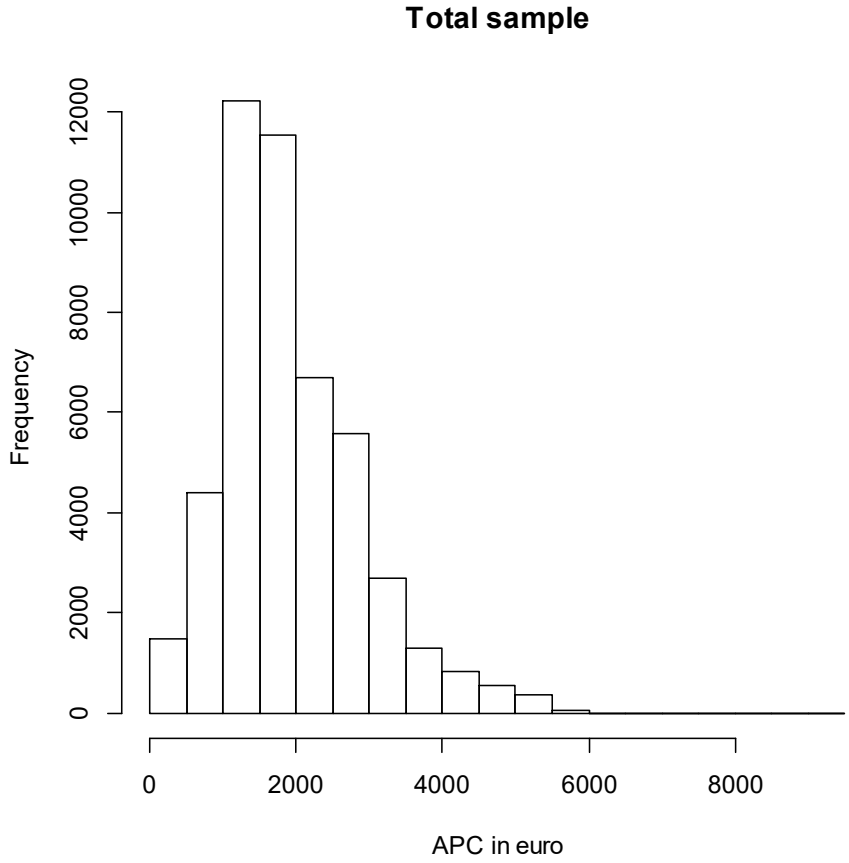
euro

```
Min.    :  40
1st Qu.:1255
Median :1738
Mean    :1924
3rd Qu.:2450
Max.    :9079
```

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Histogram of APC in euro



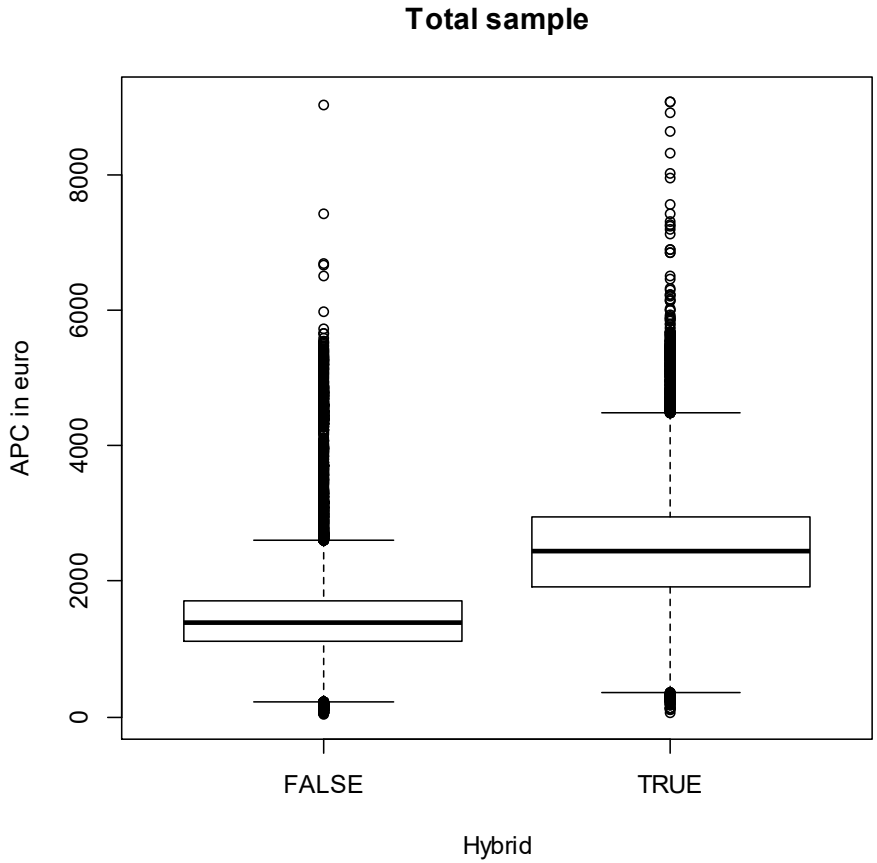
Article Processing Charges

- range mostly between 1,000 – 3,000 EUR.
- but amount sometimes to 4,000 – 6,000 EUR.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Box plots of APC in euro for OA and hybrid journals



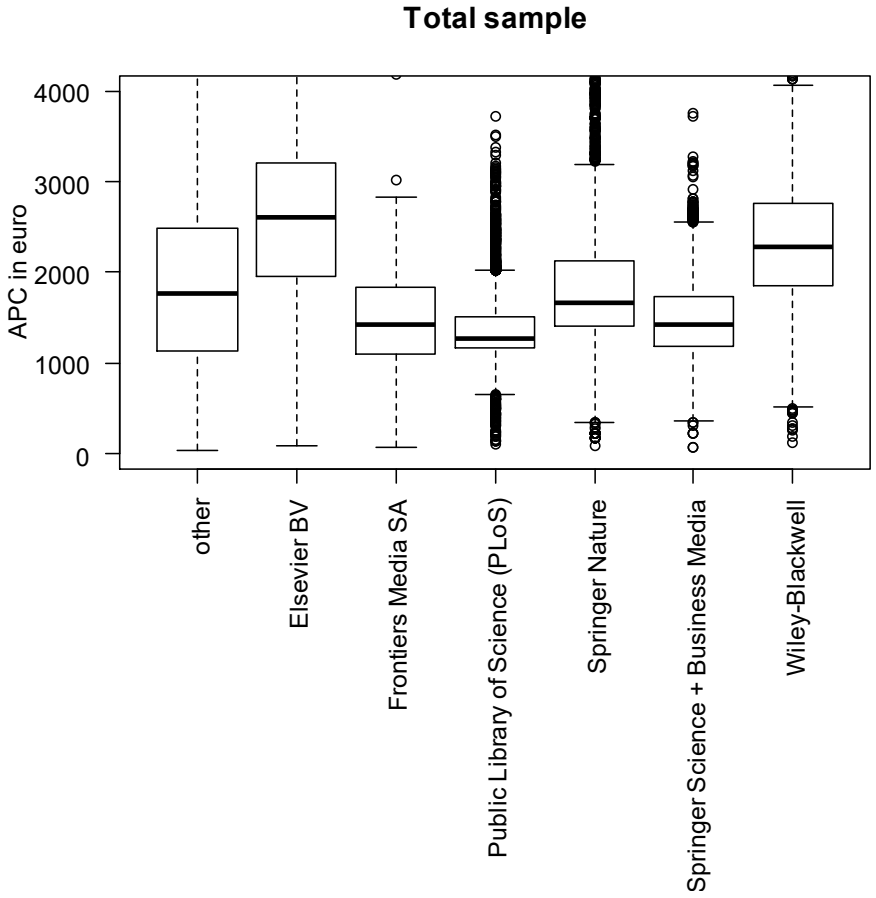
Article Processing Charges

- range mostly between 1,000 – 3,000 EUR.
- but amount sometimes to 5,000 – 6,000 EUR.
- are (on average) more expensive in hybrid journals.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Box plots of APC depending on publisher



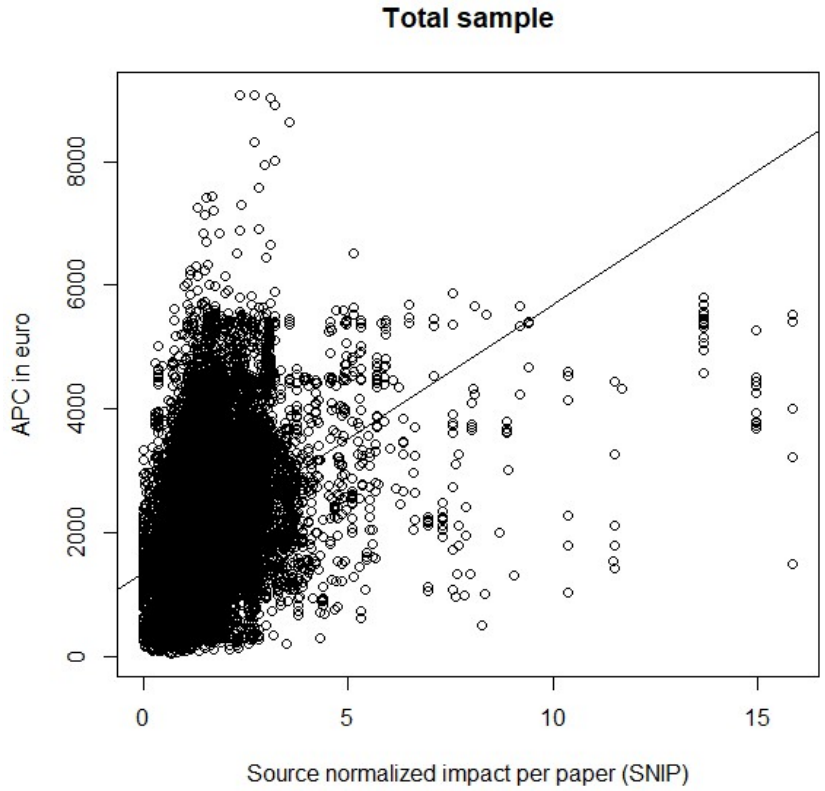
Article Processing Charges

- range mostly between 1,000 – 3,000 EUR.
- but amount sometimes to 5,000 – 6,000 EUR.
- are (on average) more expensive in hybrid journals.
- are quite different depending on publisher.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Scatter plot of APC vs. SNIP



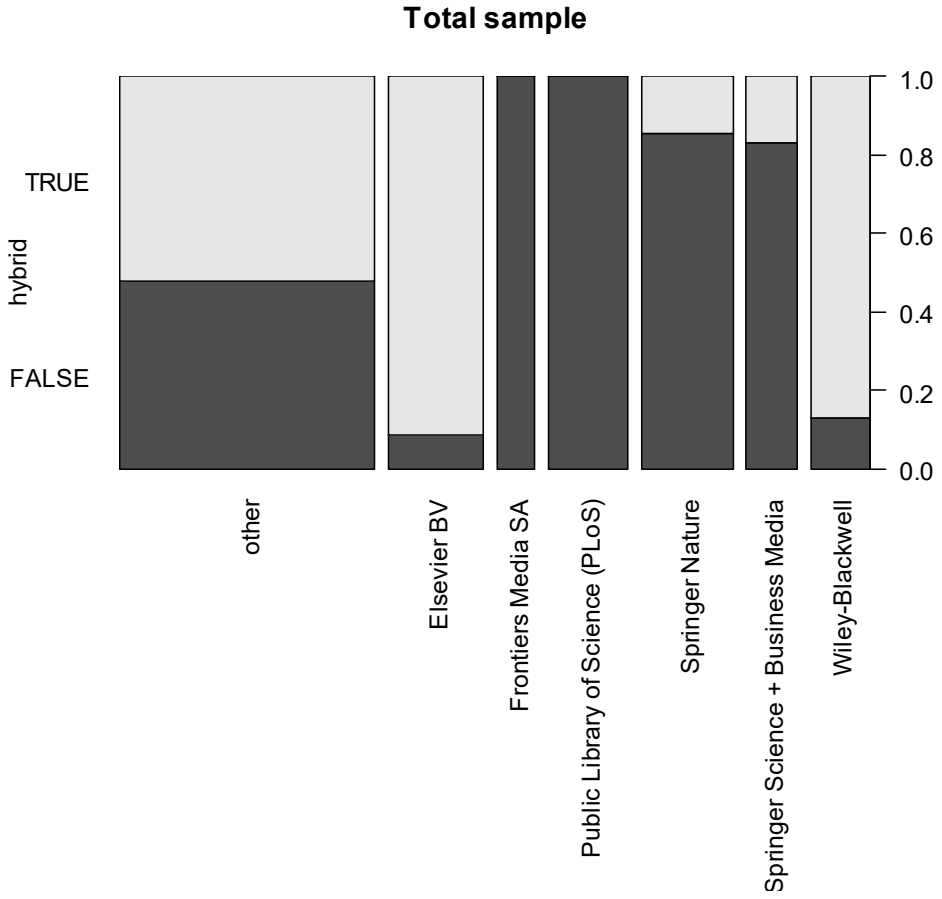
Article Processing Charges

- range mostly between 1,000 – 3,000 EUR.
- but amount sometimes to 5,000 – 6,000 EUR.
- are (on average) more expensive in hybrid journals.
- are quite different depending on publisher.
- are related to the citation impact (SNIP).

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Share of articles published in OA-/hybrid journals



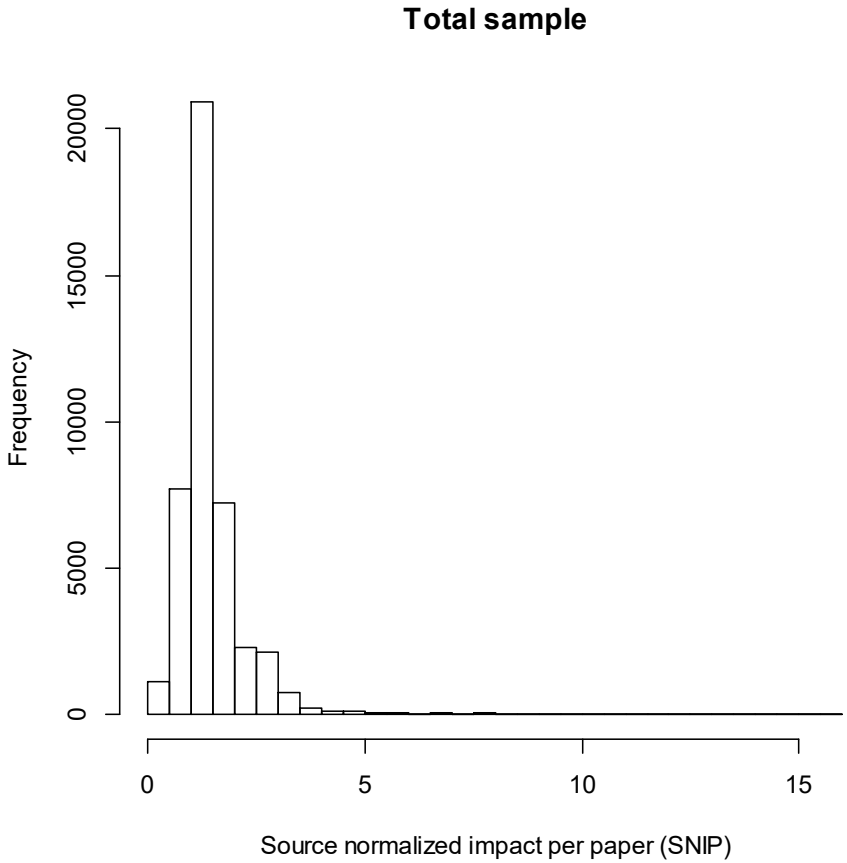
Articles are published

- often in hybrid journals at Elsevier and Wiley-Blackwell.
- often in open-access journals at Springer and Nature.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Histogram of SNIP for openAPC-records



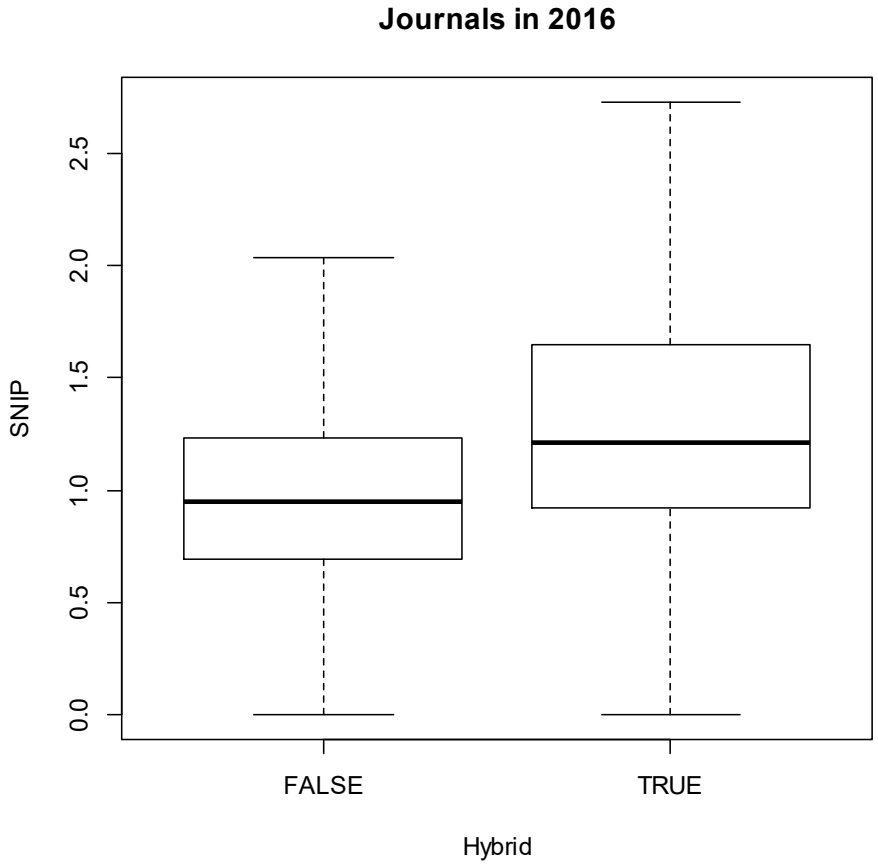
Articles are published

- often in hybrid journals at Elsevier and Wiley-Blackwell.
- often in open-access journals at Springer and Nature.
- rarely in high-impact journals.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Box plots of SNIP for open-access or hybrid journals



Journals that are

- open access tend to have lower impact.
- hybrid tend to have higher impact.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Method and statistical model

- Static linear regression with random and time effects based on *T* successive cross sections

$$APC_{it} = \alpha_i + \beta_1 SNIP_{it} + \beta_2 Hybrid_{it} + \beta_3 SNIP_{it} \times Hybrid_{it} + \mathbf{Big\_publisher}'_{it}\boldsymbol{\beta}_4 + \mathbf{Subject\_area}'_{it}\boldsymbol{\beta}_5 + \gamma_t + \epsilon_{it}$$

- Ordinary least squares (OLS)
- Heteroscedasticity-robust standard errors
- Software: R
- Sub-sample: UK, 2014–2016,
  without outliers (1%-quantile< X <99%-quantile)

Results

	Model 1	Model 2	Model 3	Model 4
(Intercept)	1797.19 (19.95)***	1800.70 (10.39)***	727.92 (40.98)***	519.38 (40.96)***
SNIP	320.42 (12.98)***		788.60 (31.82)***	728.07 (29.74)***
is_hybrid		702.61 (12.42)***	1475.81 (43.96)***	1395.93 (43.07)***
SNIP:is_hybrid			−603.29 (33.19)***	−539.69 (31.32)***
Elsevier BV				225.06 (15.76)***
Frontiers Media SA				−114.05 (31.03)***
Public Library of Science (PLoS)				−328.48 (20.28)***
Springer Nature				235.59 (22.34)***
Springer Science + Business Media				145.00 (20.60)***
Wiley-Blackwell				−29.11 (15.19)*
Life Sciences				179.48 (13.62)***
Physical Sciences				−146.77 (15.10)***
Social Sciences and Humanities				−374.95 (26.47)***
period 2015				312.13 (14.28)***
period 2016				283.40 (13.45)***
R^2	0.10	0.12	0.24	0.31
Adj. R^2	0.10	0.12	0.24	0.31
Num. obs.	22310	23818	22310	22310
RMSE	888.05	878.87	818.79	777.41

***p < 0.01, **p < 0.05, *p < 0.1

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Results

	Model 4
(Intercept)	519.38 (40.96)***
SNIP	728.07 (29.74)***
is_hybrid	1395.93 (43.07)***
SNIP:is_hybrid	−539.69 (31.32)***
Elsevier BV	225.06 (15.76)***
Frontiers Media SA	−114.05 (31.03)***
Public Library of Science (PLoS)	−328.48 (20.28)***
Springer Nature	235.59 (22.34)***
Springer Science + Business Media	145.00 (20.60)***
Wiley-Blackwell	−29.11 (15.19)*
Life Sciences	179.48 (13.62)***
Physical Sciences	−146.77 (15.10)***
Social Sciences and Humanities	−374.95 (26.47)***
period 2015	312.13 (14.28)***
period 2016	283.40 (13.45)***
R^2	0.31
Adj. R^2	0.31
Num. obs.	22310
RMSE	777.41

***p < 0.01, **p < 0.05, *p < 0.1

Base groups

- Year: 2014
- Publisher: other / smaller
- Subject area: health sciences
- Journal type: open access

Equation

- for PLoS-articles in life sciences in 2016

$$\widehat{APC} = (519 - 328 + 179 + 283) + 728 \times SNIP$$

$$\widehat{APC} = 653 + 728 \times SNIP$$

- for Elsevier hybrid-journal, else as above

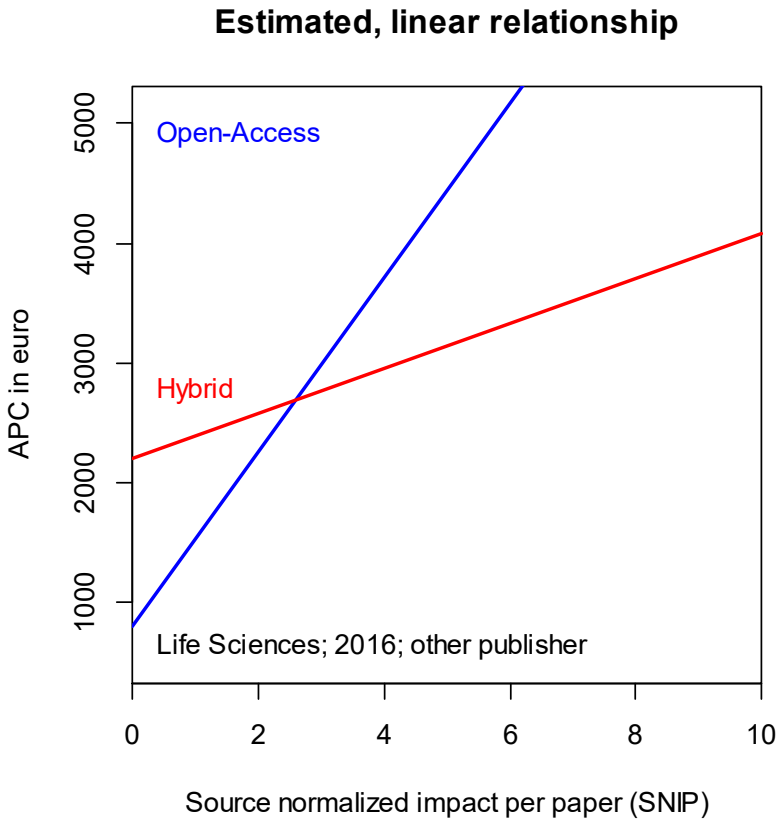
$$\widehat{APC} = (519 + 225 + 179 + 283 + 1{,}396) + (728 - 540) \times SNIP$$

$$\widehat{APC} = 2{,}602 + 188 \times SNIP$$

Results

	Model 4
(Intercept)	519.38 (40.96)***
SNIP	728.07 (29.74)***
is_hybrid	1395.93 (43.07)***
SNIP:is_hybrid	−539.69 (31.32)***
Elsevier BV	225.06 (15.76)***
Frontiers Media SA	−114.05 (31.03)***
Public Library of Science (PLoS)	−328.48 (20.28)***
Springer Nature	235.59 (22.34)***
Springer Science + Business Media	145.00 (20.60)***
Wiley-Blackwell	−29.11 (15.19)*
Life Sciences	179.48 (13.62)***
Physical Sciences	−146.77 (15.10)***
Social Sciences and Humanities	−374.95 (26.47)***
period 2015	312.13 (14.28)***
period 2016	283.40 (13.45)***
R^2	0.31
Adj. R^2	0.31
Num. obs.	22310
RMSE	777.41

***p < 0.01, **p < 0.05, *p < 0.1



OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Limitations and potential weaknesses

- Sample selection
  - openAPC possibly not a representative sample
  - due to selective recoding and reporting behavior
  - e.g. due to price cap in Germany (2,000 €)
  - ➢ Use UK sub-sample!
- Missing data
  - missing SNIP-values for 6-11% of the observations (and missing subject area)
  - related to coverage of Scopus
  - related to maturity of a journal
  - ➢ “Worst-case” analysis shows that the bias is not severe.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Conclusion

- APCs – Mirroring the impact factor?
  - In open-access journals!
  - At genuine open-access publishers!
- APCs – Legacy of the subscription-based model?
  - In hybrid journals!
  - Often at Elsevier, Springer and co.!

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Example for in-sample prediction (SNIP=1, life sciences, 2016)

- PLOS ONE article in life sciences in 2016

$$\widehat{APC} = 653 + 728 = 1{,}381\ €$$

- Article in “Journal of Neuro-science Methods” (Elsevier hybrid-journal)

$$\widehat{APC} = 2{,}602 + 188 = 2{,}790\ €$$

Equation

- for PLoS-articles in life sciences in 2016

$$\widehat{APC} = (519 - 328 + 179 + 283) + 728 \times SNIP$$

$$\widehat{APC} = 653 + 728 \times SNIP$$

- for Elsevier hybrid-journal, else as above

$$\widehat{APC} = (519 + 225 + 179 + 283 + 1{,}396) + (728 - 540) \times SNIP$$

$$\widehat{APC} = 2{,}602 + 188 \times SNIP$$

In-sample APC-prediction

	PLoS, OA	Elsevier, hybrid
SNIP=1	$\widehat{APC}_{it} = \text{EUR } 1381$	$\widehat{APC}_{it} = \text{EUR } 2790$
SNIP=1.37	$\widehat{APC}_{it} = \text{EUR } 1650$	$\widehat{APC}_{it} = \text{EUR } 2860$
SNIP=1.81	$\widehat{APC}_{it} = \text{EUR } 1971$	$\widehat{APC}_{it} = \text{EUR } 2942$
SNIP=2	$\widehat{APC}_{it} = \text{EUR } 2109$	$\widehat{APC}_{it} = \text{EUR } 2978$
SNIP=15	$\widehat{APC}_{it} = \text{EUR } 11573$	$\widehat{APC}_{it} = \text{EUR } 5422$

Note: The in-sample APC prediction for an open-access journal with a SNIP-score of 15 is are rather hypothetical consideration, as no open-access journal has comparable impact.

Actual and predicted total amount of APCs

	Total amount of APCs, in euro
UK, actually paid	52,658,541
UK, as if all OA	44,662,308
UK, as if all hybrid	56,863,847
Total, actually paid	83,969,558
Total, as if all OA	72,229,822
Total, as if all hybrid	101,031,495

Note: Only complete cases.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

PredictedAPCs.html

Predicted APCs (out-of-sample, preliminary results)

Predicted APCs

journal_full_title	SNIP	Subject.area	P	pred.APC.OA	pred.total.APC.OA	pred.APC.hyb	pred.total.APC.hyb
All	All	All	All	All	All	All	All
Ca-A Cancer Journal for Clinicians	88.16	Health Sciences	73	69508	1691361	19619	477396
MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports / Centers for Disease Control	32.53	Physical Sciences	16	25954	138421	8519	45435
MMWR. Surveillance summaries : Morbidity and mortality weekly report. Surveillance summaries / CDC	19.65	Physical Sciences	31	15913	164434	5991	61907
Vital and health statistics. Series 10, Data from the National Health Survey	17.43	Physical Sciences	3	14182	14182	5555	5555
Reviews of Modern Physics	15.29	Physical Sciences	113	12513	471323	5135	193418
The Lancet	14.8	Health Sciences	1015	12314	4166237	5222	1766777
New England Journal of Medicine	13.4	Health Sciences	1526	11223	5708766	4948	2516883
Living Reviews in European Governance	12.27	Social Sciences & Humanities	2	9923	6615	4307	2871

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Show 100 entries

Search:

				Predicted APCs			
journal_full_title	SNIP	Subject.area	P	pred.APC.OA	pred.total.APC.OA	pred.APC.hyb	pred.total.APC.hyb
All	All	All	All	All	All	All	All
ACS Sensors	1.24		203				
Journal of the Textile Institute	1.23	Life Sciences	467	1849	287828	2674	416253
Biological Control	1.23	Life Sciences	486	1849	299538	2674	433188
Animal Behaviour	1.23	Life Sciences	924	1849	569492	2674	823592
Asian-Australasian Journal of Animal Sciences	1.23	Life Sciences	672	1849	414176	2674	598976
Aquacultural Engineering	1.23	Life Sciences	156	1849	96148	2674	139048
Aquatic Toxicology	1.23	Life Sciences	860	1849	530047	2674	766547
Marine Pollution Bulletin	1.23	Life Sciences	2015	1849	1241912	2674	1796037

Showing 4,001 to 4,100 of 22,948 entries

Previous 1 … 40 41 42 … 230 Next

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Show 100 entries

Search:

				Predicted APCs			
journal_full_title	SNIP	Subject.area	P	pred.APC.OA	pred.total.APC.OA	pred.APC.hyb	pred.total.APC.hyb
All	All	All	All	All	All	All	All
Pharmaceutical Journal	0.05	Life Sciences	118	929	36541	2442	96052
Journal of International Pharmaceutical Research	0.05	Life Sciences	559	929	173104	2442	455026
Klinicka Farmakologie a Farmacie	0.05	Life Sciences	109	929	33754	2442	88726
Suvremena Psihologija	0.05	Social Sciences & Humanities	35	396	4620	1909	22272
ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant	0.05	Social Sciences & Humanities	159	396	20988	1909	101177
Journal of Third World Studies	0.05	Social Sciences & Humanities	44	396	5808	1909	27999
Carrefours de l'Education	0.05	Social Sciences & Humanities	76	396	10032	1909	48361
African Journal of Legal Studies	0.05	Social Sciences & Humanities	48	396	6336	1909	30544

Showing 21,201 to 21,300 of 22,948 entries

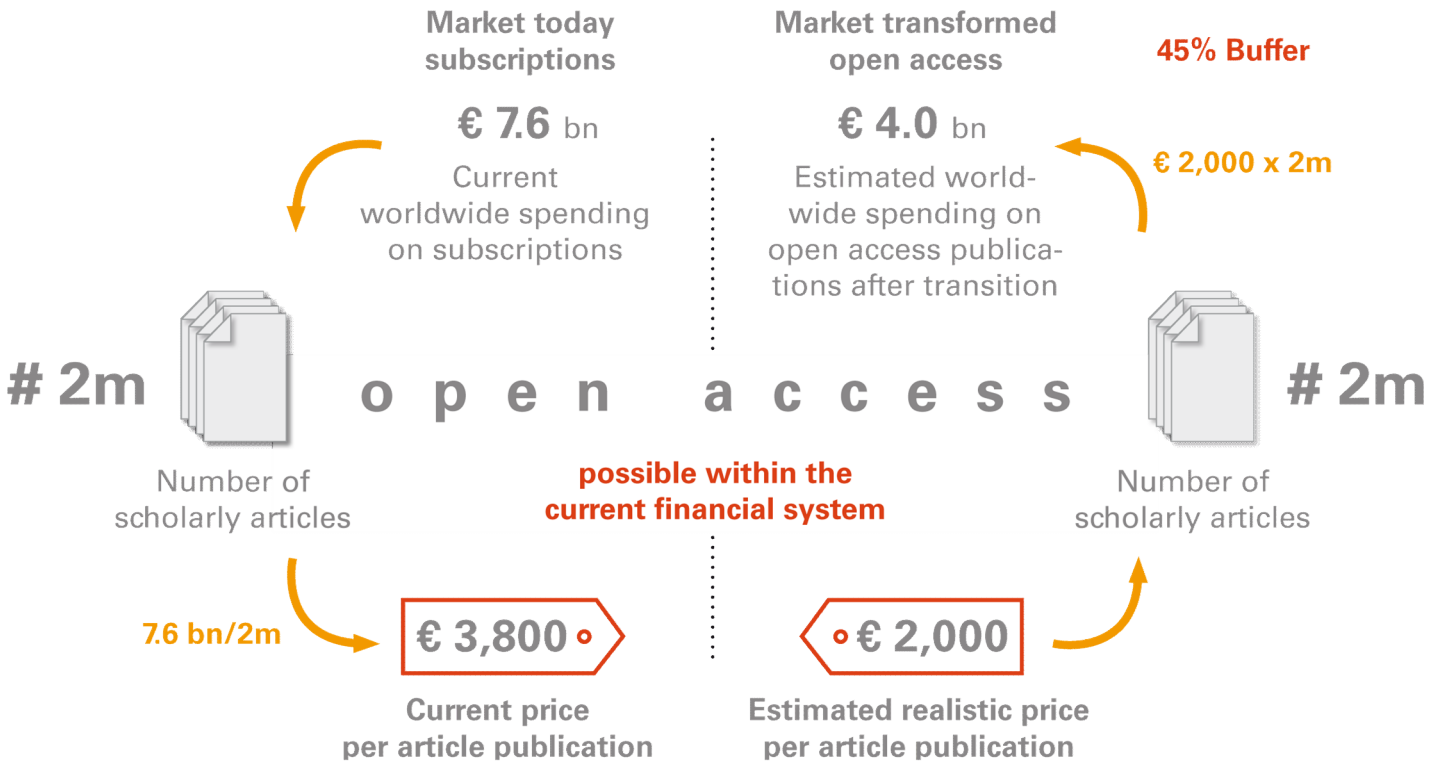
Previous 1 … 212 213 214 … 230 Next

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

The global perspective (source: MPDL Whitepaper)

Worldwide Publishing Market

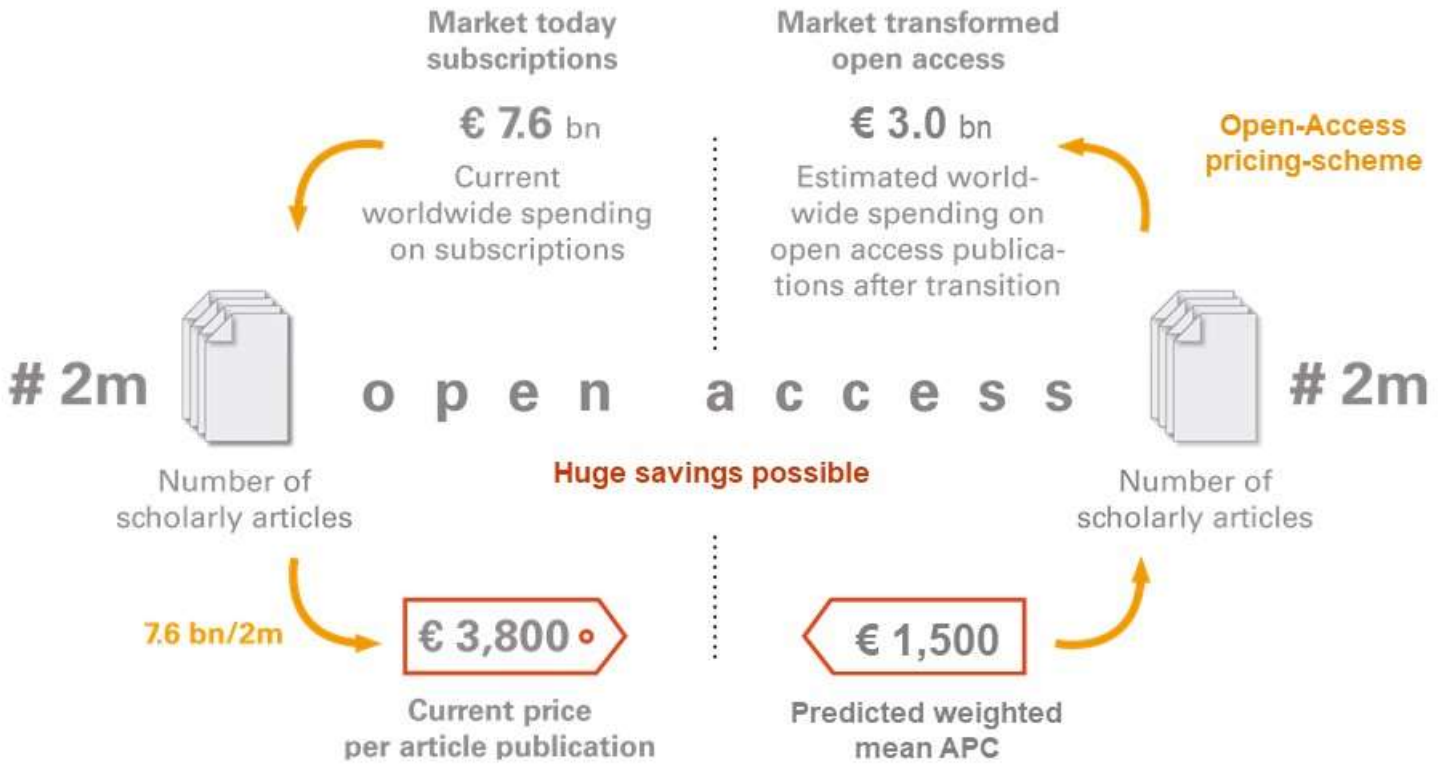


OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

The global perspective in style of the MPDL Whitepaper

Worldwide Publishing Market

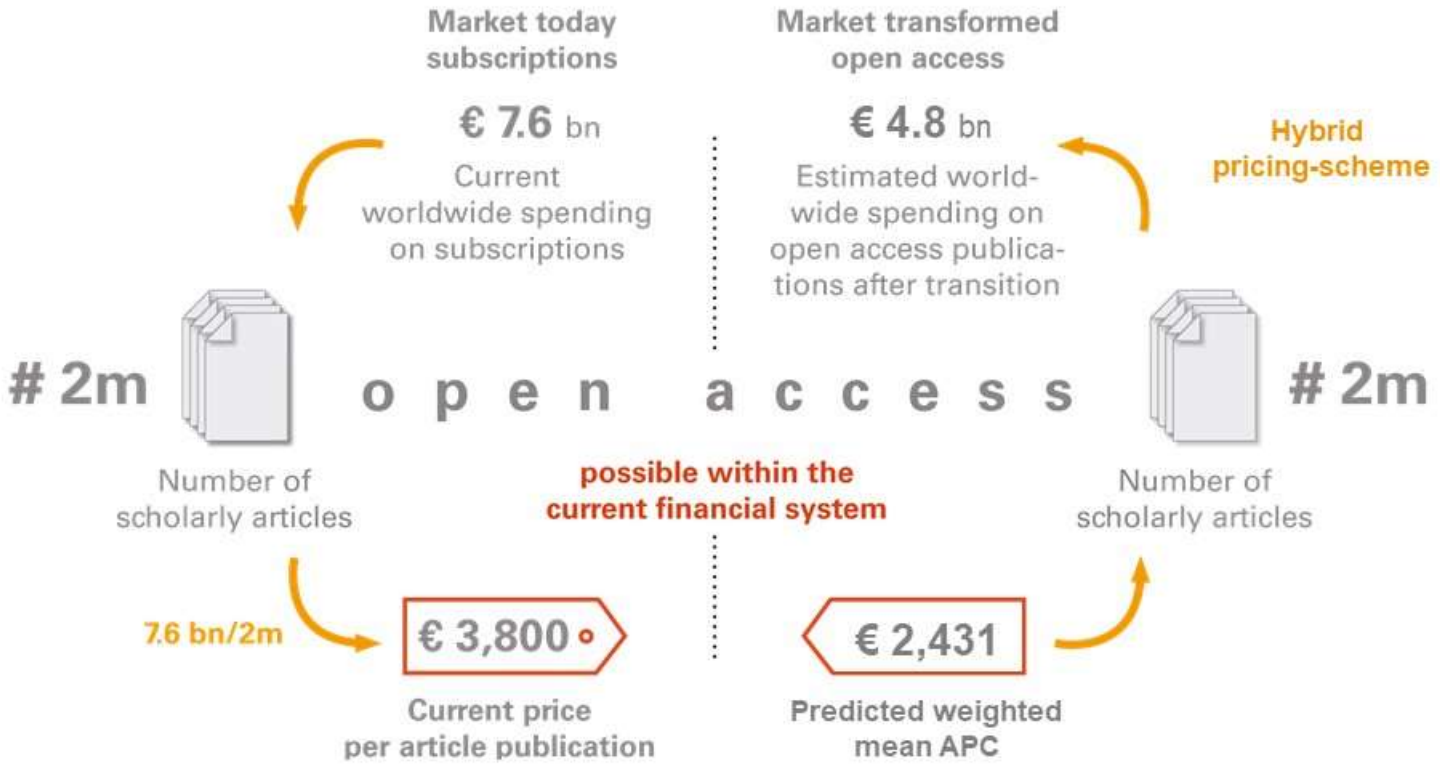


OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

The global perspective in style of the MPDL Whitepaper

Worldwide Publishing Market



OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

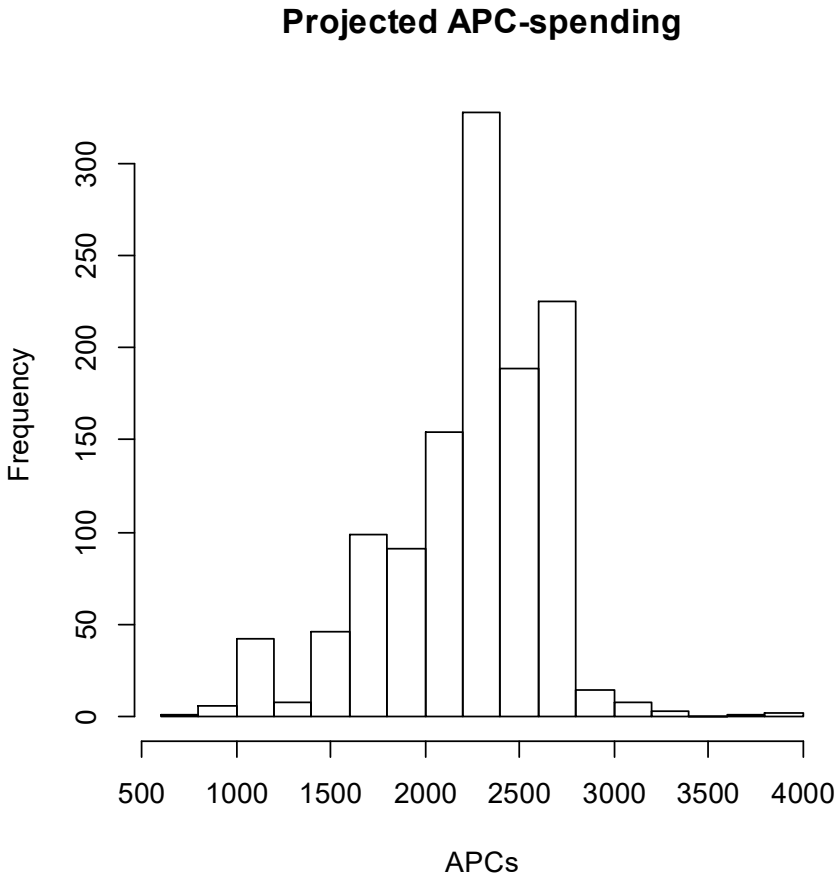
Publication output in Bielefeld (Source: Web of Science)

	2015	2016	2017	Percent with funding information
Publication with affiliation	1109	981	1019	62%
– Articles in journals	915	741	856	74%
– Papers in conference proceedings	89	109	60	27%
– other	105	131	103	6%
Publications with reprint author	557	453	502	61%
– Articles in journals	468	335	415	70%
– Papers in conference proceedings	49	78	47	33%
– other	40	40	40	8%

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Projected APC-spending in Bielefeld



	All articles	Articles without funding information
Total APC-spending p.a.	906,581	274,488
Average APC	2,235	2,244
Median APC	2,324	2,299
Need on external funds	0	632,093
No. of not included articles	1	0
Break-even APC		
Break-even APC incl. proceeding papers		

Note: APCs in euro; preliminary results. Only publications are considered where the reprint author is affiliated with Bielefeld.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Questions?

Dr. Nina Schönfelder
National Contact-Point Open-Access OA2020-DE
Bielefeld University Library
Universitätsstr. 25 | D-33615 Bielefeld
Tel.: +49 (0) 521/106-2546 | email: nina.schoenfelder@uni-bielefeld.de

www.oa2020-de.org
@oa2020de