OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

APCs
Mirroring the impact factor or legacy of the subscription-based model?

UKSG Annual Conference
Telford, 08–10.04.2019

Dr. Nina Schönfelder

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

National Contact Point Open Access OA2020-DE

- **Project goal:**
  Create prerequisites for the large-scale open access transformation in Germany
- **Working Package 4:**
  Analysing financial flows, shaping financial models, and consultation with funders

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Steps

- Focus here: **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

- Follow-up study 1: Projecting APCs for currently hybrid or closed-access journals
- Follow-up study 2: Comparing projected total APC-spending with libraries budgets‘ after a hypothetical full journal flipping for each German university and research institute

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Agenda

- Data
- Method
- Results
- Conclusion

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Data

- OpenAPC data set (from 2018-02-06)
  - Aggregates data from Jisc yearly collections of cost data and the Wellcome Trust publishing yearly reports of all their funded articles, among others
  - 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

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                      : 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
PLoS                 : 5690
Wiley-Blackwell      : 4265
Springer S+B Media   : 3627
Frontiers Media SA   : 2718
(Other)              :18120
```

SNIP

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

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
```

Subject.area

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

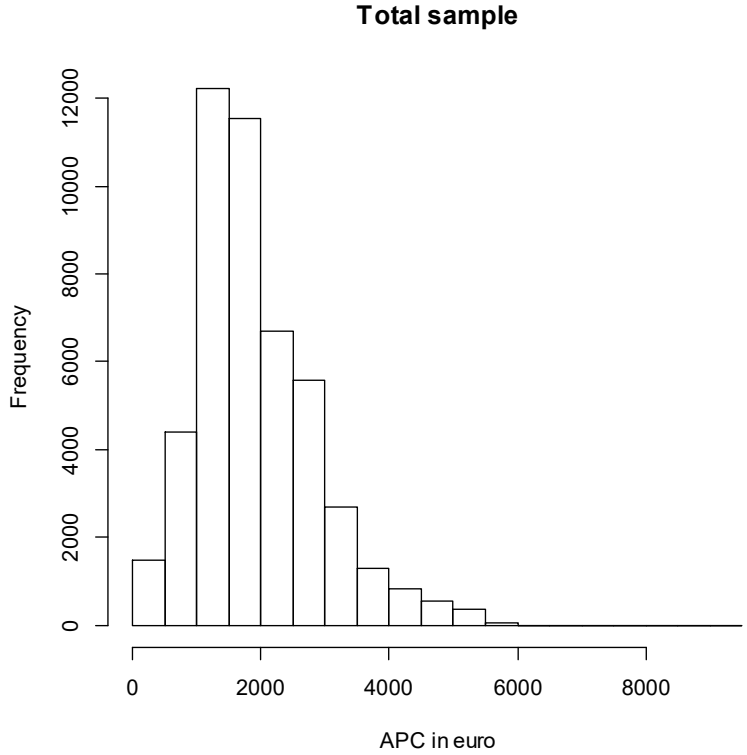
euro

```
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



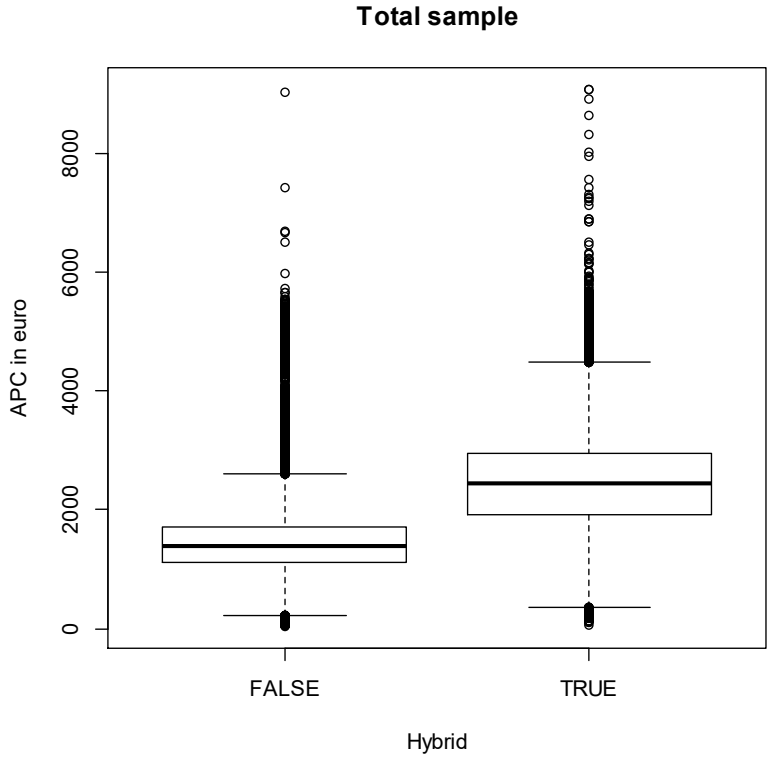
Article Processing Charges

- range mostly between 1,000 – 3,000 EUR.
- but amount sometimes to 5,000 – 6,000 EUR.

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

Box plots of APC in euro for open-access and hybrid journals



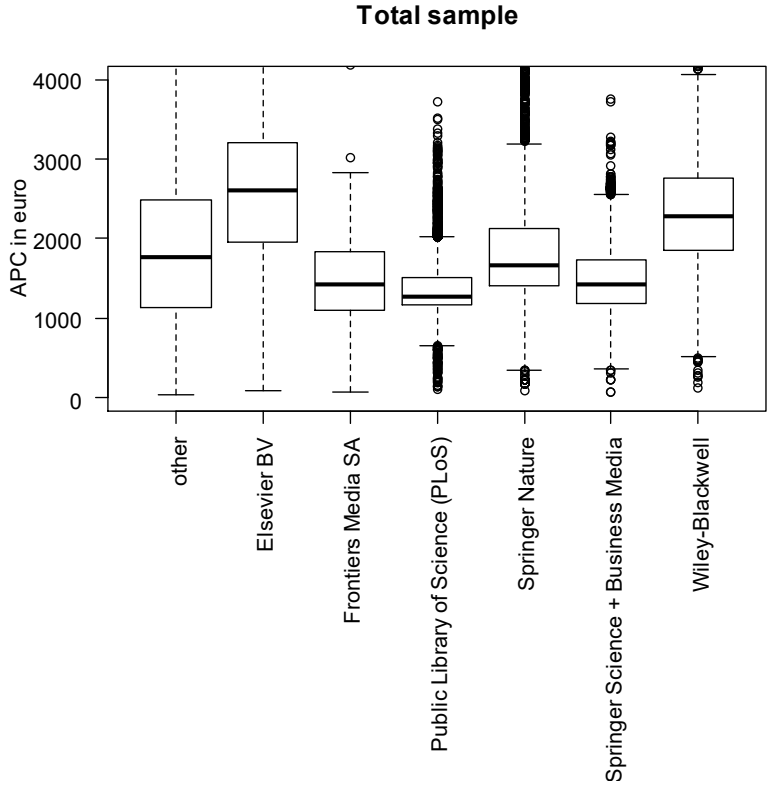
Article Processing Charges

- 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



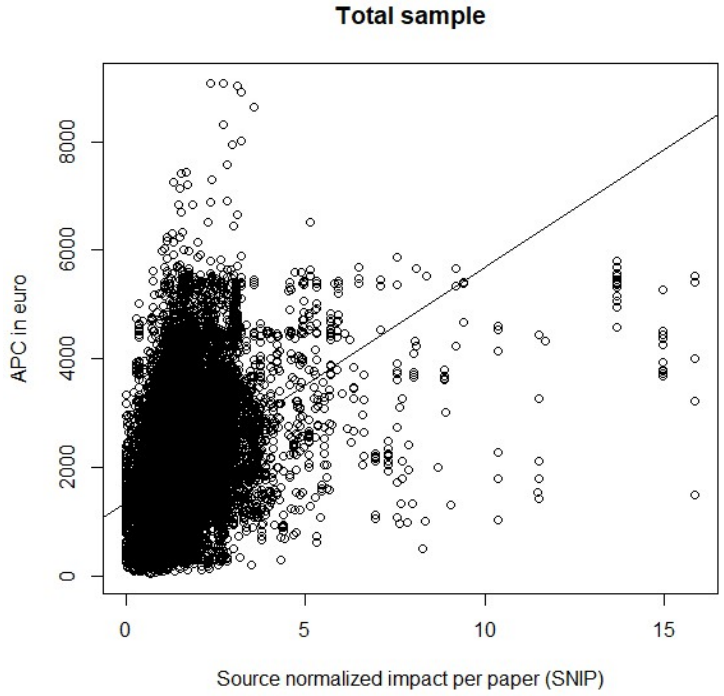
Article Processing Charges

- 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



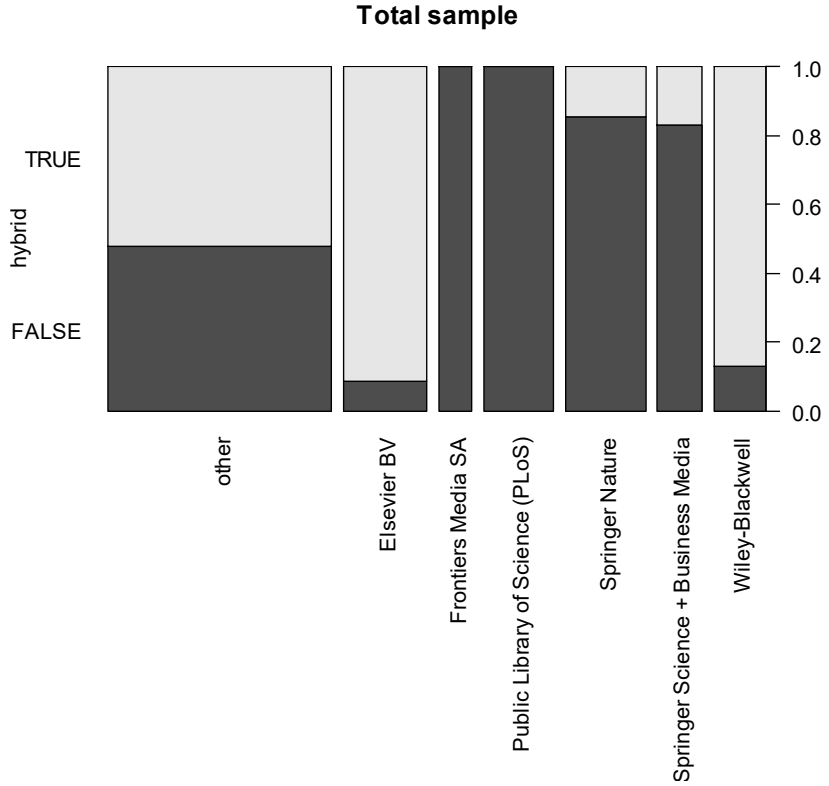
Article Processing Charges

- 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 open-access vs. hybrid journals



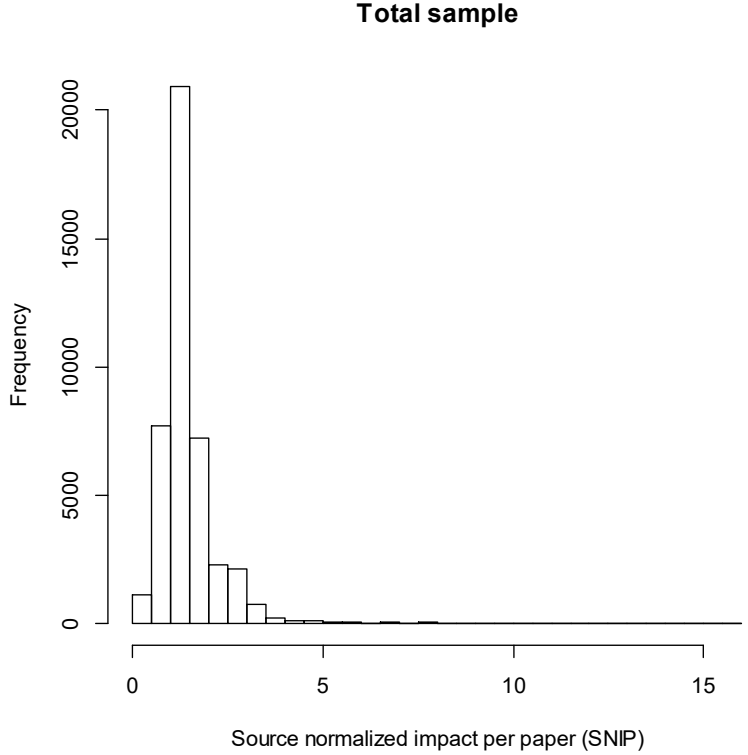
Articles are published

- 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



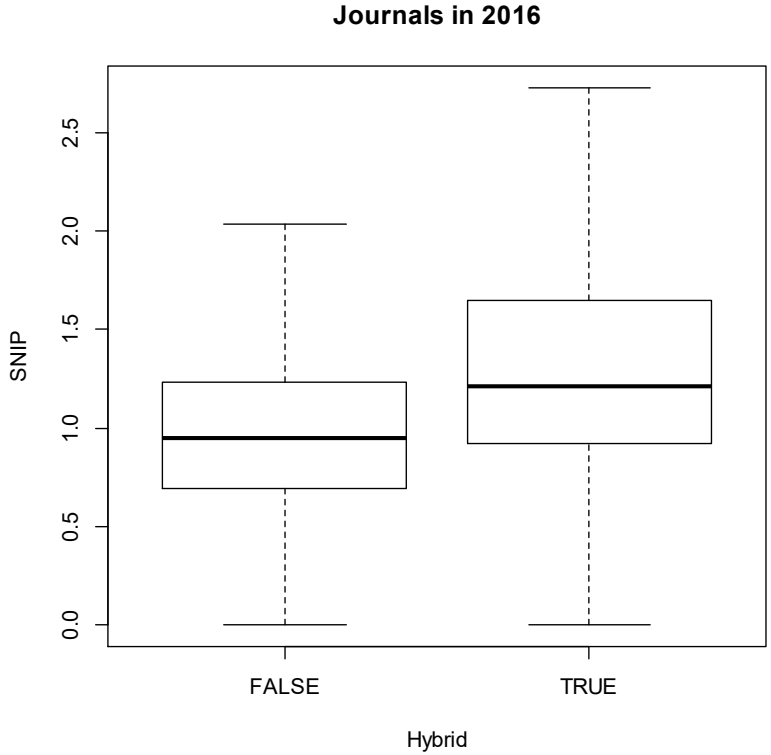
Articles are published

- 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

- Multivariate linear regression of

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

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

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

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 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 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

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

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

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

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 Neuroscience 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

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

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

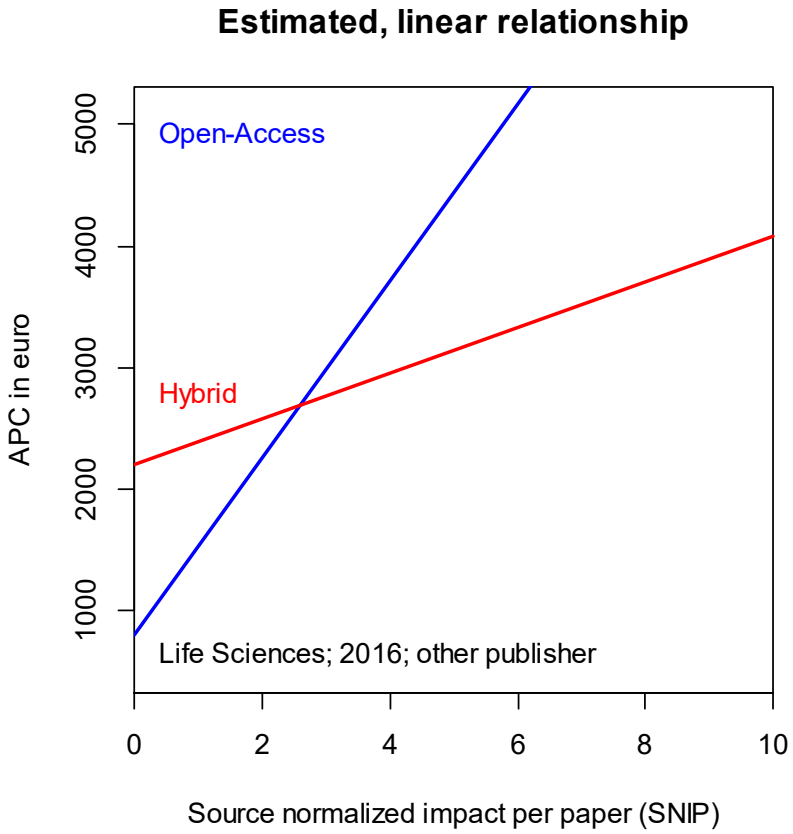
OPEN ACCESS 2020 DE

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



OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

In-sample APC-prediction

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

OPEN ACCESS 2020 DE

Nationaler Open-Access-Kontaktpunkt

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

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

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