

Investigating the blind spot of monitoring systems for article processing charges in full OA journals: An approach for the estimation of publication fees not covered by OpenAPC¹

Andre Bruns (<https://orcid.org/0000-0002-2976-0826>)

Niels Taubert (<https://orcid.org/0000-0002-2357-2648>)

Introduction

In recent years, a number of activities could be observed that aim to support the transformation towards Gold Open Access (OA), which is based on article processing charges (APC). On the organizational level, a large number of research institutions created central funds to cover publication fees for OA publications of their authors (in short: publication funds) and established structures and workflows for the organization of payments. On the level of countries, nation-wide OA-contracts have been negotiated, making a bulk of the publication output of those countries OA. Examples can be found in Austria, Finland, Hungary, Germany, the Netherlands, Norway, Poland, Qatar, Sweden and the UK whereby institutions combine spending on subscriptions and OA to shift the balance towards OA publication as the standard.

An important advantage of the APC-based OA publishing is that more transparency is possible regarding the expenditures and financial flows for publications. In the subscription model, details of the licenses are kept secret as subscription contracts often contain nondisclosure agreements. Therefore, contracts between research organisations and publishers can hardly be compared. This situation has changed with the introduction of monitoring instruments for APC. Such systems cover data of actual APC-payments, which allows us to deepen our understanding of the OA transformation and provides important information for future planning. However, the value of monitoring instruments does not only depend on the creation of standardized procedures and reporting routines for quality controlled and comparable data but also on the size and completeness of the data covered by them. An ideal APC monitoring instrument would cover complete APC payments from all research organizations of a given domain. In the real world, APC monitors are lacking for at least two reasons: first, not all institutions in a given domain deliver data to APC monitors, mostly because of the fact that not all of them have a central publication fund that processes APC payments and collects the data of these transactions. Second, even research institutions that have a central publication fund usually do not process all payments via this fund. A smaller or larger number of APC payments are made by different entities of a research organization, are processed in various ways, and can therefore not easily be captured by monitoring instruments. Some research organizations tried to catch APC payments more exhaustively and extracted all payments to publishers from the central administration. Examples are the German Forschungszentrum Jülich (Barbers et al. 2018) and Stockholm University (Loven 2019).

The current article focusses on the world's largest collection of APC payments, the OpenAPC dataset (Pieper & Broschinski 2018)², and addresses parts of this desideratum. It compares

¹ This work was funded by the Federal Ministry of Education and Research (BMBF, FKZ: 160A32).

² <https://openapc.net/> (accessed May 5th 2021).

two methods that aim to estimate the entire spending for APC in so-called full OA journals, i.e. journals making their whole content freely available online without delay. The first method is simpler and is based on the comparison of the volume of different subsets of the publication output of research institutions. However, it tackles the problem of incomplete coverage of institutions in OpenAPC only and might be less precise. The second method addresses both shortcomings (incomplete coverage of payments of participating institutions as well as incompleteness of institutions) and is based on the identification of probable APC-liable publications. The goal of both model calculations is to contribute to a more realistic picture of the amount of money that is currently spent for Gold OA publications.

The article is organized as follows: in a first step, the literature about OA- and APC-monitoring is reviewed. This is followed by a more detailed description of the guiding question in the second step. In a third step, the two estimation methods are explained. Step four describes the results of the estimation, followed by a discussion about the strength and limitations of the two approaches in a fifth step.

1. Literature Review

This article contributes to a growing body of research that studies the characteristics of a transformation towards OA based on APC. For an appropriate understanding of the character and the current state of the OA transformation, it is important to note that there are various types of OA and that the APC-based full OA is only one of them. Still, the most important types are

- *Green OA*, i.e. pre- and postprints that are available via institutional and disciplinary/subject repositories (Guédon 2004: 315, Suber 2012: 5),
- *Moving wall or delayed OA*, i.e. publications made openly available online by the publisher after an embargo period (Willinsky 2003, Laakso & Björk 2013)
- *Hybrid OA*, i.e. OA provided by subscription-based journals that allow authors to make their individual article immediately available online if article processing charges have been paid (Prosser 2003, Björk 2012, Laakso & Björk 2016: 920) and
- *Full OA*, i.e. publications in cover-to-cover openly accessible journals and conference proceedings, allowing immediate access at the time of publication (Carroll 2011).³

APC are applied in the context of two OA types, hybrid and full OA. In the context of the hybrid model, APC occur only in a fraction of the journal publication output. It is therefore not possible to automatically assume that an APC payment has been made if a publication appears in a hybrid journal. In the case of Full OA, one has to note that a number of journals within this category do not charge APC. Such journals are called platinum (Wilson 2007) or diamond OA (Fuchs & Sandoval 2013). At the global level, roughly two thirds of the journals are included in the Directory of Open Access Journals (DOAJ)⁴ (Morrison et al. 2015). The application of APC seems to differ by field (Crawford 2017). For example, for medicine, two thirds of the journals

³ For the contribution of the different OA types to the overall share of publications that are freely accessible online, see Laakso et al. 2011, Gargourie et al. 2012, Archambault et al. 2014, Crawford 2015, 2017 et al., Wohlgemuth et al. 2017, Piwowar et al. 2018, Martín-Martín et al. 2018, Abediyarandi & Mayr 2019, Hobert et al. 2020.

⁴ <https://doaj.org/> (accessed on May 5th 2021).

do not impose APC (Asai 2019). In addition, the take-up of APC also varies by region. A large share of OA journals not charging APC can be found in Latin America (Appel & Albagli 2019), the Middle East, and Eastern Europe (Crawford 2017). They are financed by other means, such as subsidies from the state as in the case of Brazilⁱ, grants and support from learned societies, or they are driven by the voluntary and unpaid work of dedicated scientists.

A third set of studies that is relevant in this context analyzes the price for publishing in an APC environment. Because of the lack of other data, early studies referred to list prices on publishers' websites (Morrison et al 2015) or to prices as recorded by DOAJ (Björk & Solomon 2015). Given that the amount of money that is actually paid for APC differs from such list prices, and given that payments for articles published in the same journal may also vary, more recent studies draw on collections of actual payments since such data collections are now available (Jahn & Tullney 2016). Regarding average prices paid for APC, the reported numbers vary at a similar scale between €905 (Asai 2019) and €1,479 (Pieper & Broschinski 2018). One peculiarity is that all studies report large standard deviations, indicating that there is much variance in the pricing of APC of publishers. A second key aspect in the analysis of the price structure of APC concerns the determinants for the price. There is some evidence that APC prices

- are higher for publications in hybrid than in full OA journals (Pinfield et al. 2016; Jahn & Tullney 2016, Schönfelder 2020),
- vary by discipline (Solomon & Björk 2012),
- vary by the type of publisher (Asai 2019),
- vary by impact as measured by average number of citation indicators like the journal impact factor (JIF) or Source Normalized Impact per Paper (SNIP) (Solomon & Björk 2015; Schönfelder 2020),
- vary by the language of the journal (Asai 2019).

2. Research Question

Monitoring systems for article processing charges are a valuable instrument for understanding and further planning of a transformation towards OA based on APC. The aim of the article is to develop, compare and evaluate two methods for an estimation of the expenses for APC of universities that do not contribute to monitoring systems. The first is based on the proportion of different subsets of the publication output of universities covered by monitoring systems. It estimates the proportion of APC-liable publications not covered by them. The second is based on the identification of APC-liable publications in the publication output of universities not covered by APC monitoring systems. Given that the publication output can be determined for German universities, this country is taken as an example for this study. For practical reasons, the article is limited to publications in full OA journals.

3. Methodology and data

The study exploits the following three data sources:

- *Web of Science (Wos)*: In a first step, the publication output was determined for all German universities. Given that there is no database that exhaustively covers their entire publication output, we used the Web of Science (WoS) database hosted by the competence centre for bibliometrics in its version of 2019.⁵ Although WoS is known for a selective coverage of the publication output of countries and institutions and for various biases (Mongeon and Paul-Hus 2016), the advantage of this version of the database is that it is enriched with disambiguated institutional addresses for German institutions (Winterhager et al. 2014, Rimmert et al 2017). This allows us to precisely identify the publication output of research institutions in that source. An exhaustive list of German universities was compiled and all author-address-combinations for publications with at least one address from a German university were retrieved from the database. This information also includes the identifier of the institution, corresponding author information, first author information publication identifier (DOI and WoS-Identifer), article title, publication year, publication type, number of authors, and identifiers of the serial (ISSN, P-ISSN, E-ISSN, ISSN-L). Information whether or not the university contributes to OpenAPC was added. Since the study is interested in an estimation of APC payments, and the institution of the corresponding author is usually supposed to cover the costs, a publication is only attributed to the university of the corresponding author.
- *ISSN-Gold-OA-list*: in a second step, publications in full OA journals were identified for the entire publication output of German universities covered by WoS. The ISSN-Gold-OA-List was used in the Version 4.0 as a source of evidence for publications in full OA journals (Bruns et al. 2020). It aggregates different full OA journal lists, including the Directory of Open Access Journals (DOAJ), PubMedCentral (PMC), Directory of Open Access Scholarly Resources (ROAD) and OpenAPC.
- *OpenAPC*: in a third step, confirmed payments in the period 2017-2019 for publications of German universities were harvested from OpenAPC on August 28th 2020. OpenAPC include publications with APC payments from universities that hold a central publication fund. Nevertheless, it is considered incomplete as payments may have been processed outside the publication funds.

The three data sources and their coverage can be defined and illustrated (see Figure 1) as follows:

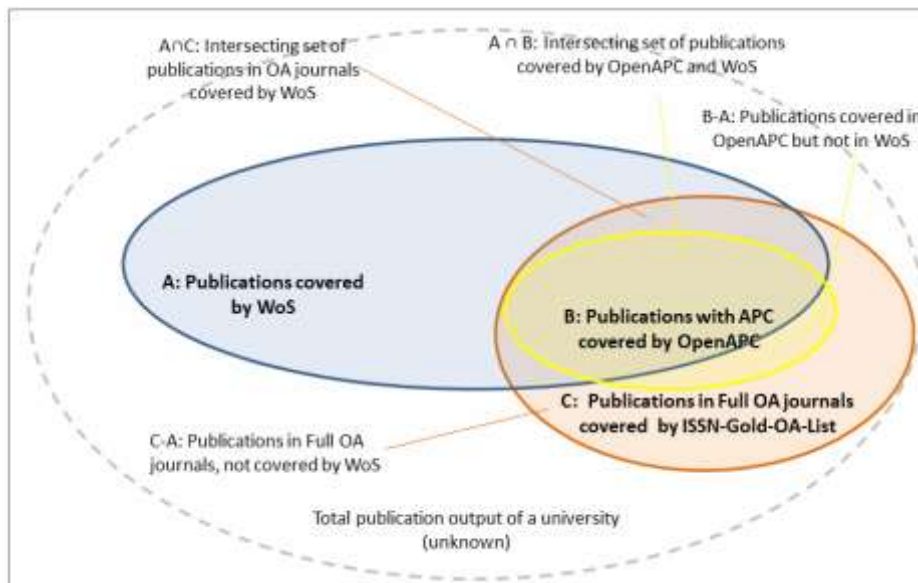
Definitions

- A publications with a corresponding author of a certain university (in what follows ‘corresponding author publication’) covered by WoS
- B corresponding author publications with APC as documented by OpenAPC (OpenAPC universities only)
- C corresponding author publications of a certain university in full OA journals as documented by ISSN-Gold-OA-list 4.0 ($B \subset C$)
- $A \cap B$ corresponding author publications of a certain university covered by WoS and OpenAPC
- $A \cap C$ corresponding author publications of a certain university covered by WoS and ISSN-Gold-OA-list 4.0 ($A \cap B$ is part of $A \cap C$)
- $B \setminus A$ corresponding author publications of a certain university covered by OpenAPC but not covered by WoS

⁵ <http://www.forschungsinfo.de/Bibliometrie/en/index.php?id=home> (accessed on May 5th 2021).

$C \setminus A$ corresponding author publications of a certain university covered by ISSN Gold OA list 4.0 but not covered by WoS

Figure 1: Sets of publications in WoS, OpenAPC, and ISSN-Gold-OA-list



All German universities were assigned to one of two exclusive groups. The first group (in what follows ‘OpenAPC universities’) includes all universities with payment information on OpenAPC for all years of the period 2017-2019. It consists of 41 universities. The second group includes all other German universities that have at least one corresponding author publication in 2019 and that do not provide data to OpenAPC. In what follows, it is called ‘non-OpenAPC universities’.

Based on the three data sources and on the different intersecting sets of publications, two estimations of the expenditures for APC were calculated for universities that do not contribute to APC.

3.1 Global estimation

The first procedure ‘global estimation’ is based on the assumption that the proportions of the different sets of publications are similar for OpenAPC universities and non-OpenAPC universities. In addition, it is assumed that the average costs for APC are similar for both OpenAPC and non-OpenAPC universities.

The procedure is organized in three steps.

For each OpenAPC university, the three basic sets of publications A, B, C as well as the intersecting sets $A \cap B$ and $A \cap C$ and $B \setminus A$ are determined.⁶ In addition, the proportions of different sets of publications are calculated.

⁶ Please note that $C \setminus A$ cannot be determined as the calculation would require a data base that covers all publications of a certain university in journals of the ISSN-Gold-OA-list. As far as we know, such a data base does not exist.

In a second step, B is estimated for all non-OpenAPC universities by exploiting the information about the proportions of the different sets of publications of OpenAPC universities. Two alternative estimations are possible.

The first estimation uses information about the proportions of A (WoS covered corresponding author publications) and B (OpenAPC covered publications) only.

$$B \sim_{(N_OAPC\ U)} = A \cap B \sim_{(N_OAPC\ U)} + B \setminus A \sim_{(N_OAPC\ U)}$$

where

$$A \cap B \sim_{(NOAPC\ U)} = \left(\frac{\sum (A \cap B_{(OAPC1)} + A \cap B_{(OAPC2)} + [\dots] + A \cap B_{(OAPCn)})}{\sum (A_{(OAPC1)} + A_{(OAPC2)} + [\dots] + A_{(OAPCn)})} \right) * A_{(N_OAPC\ U)}$$

and

$$B \setminus A \sim_{(NOAPC\ U)} = \left(\frac{\sum (B \setminus A_{(OAPC1)} + B \setminus A_{(OAPC2)} + [\dots] + B \setminus A_{(OAPCn)})}{\sum (A \cap B_{(OAPC1)} + A \cap B_{(OAPC2)} + [\dots] + A \cap B_{(OAPCn)})} \right) * A \cap B \sim_{(NOAPC\ U)}$$

The second estimation is a variation of the first one. It estimates $B \setminus A$ with the same method but differs regarding the estimation of $A \cap B$ as it includes information about the set of publications in Full OA journals. Even though the estimation is more complex, it would be more precise if the proportions of the sets of publications B and C vary less than the proportions of A and B. In other words, this would be the case if the proportion of APC-liable publications of Gold OA publications varies less than the proportion of APC-liable publications in the publication output of a university.

$$B \sim_{(NOAPC\ U)} = A \cap B \sim_{(NOAPC\ U)} + B \setminus A \sim_{(NOAPC\ U)}$$

where

$$A \cap B \sim_{(NOAPC\ U)} = \left(\frac{\sum (A \cap B_{(OAPC1)} + A \cap B_{(OAPC2)} + [\dots] + A \cap B_{(OAPCn)})}{\sum (A \cap C_{(OAPC1)} + A \cap C_{(OAPC2)} + [\dots] + A \cap C_{(OAPCn)})} \right) * A \cap C_{(NOAPC\ U)}$$

and

$$B \setminus A \sim_{(NOAPC\ U)} = \left(\frac{\sum (B \setminus A_{(OAPC1)} + B \setminus A_{(OAPC2)} + [\dots] + B \setminus A_{(OAPCn)})}{\sum (A \cap B_{(OAPC1)} + A \cap B_{(OAPC2)} + [\dots] + A \cap B_{(OAPCn)})} \right) * A \cap B \sim_{(NOAPC\ U)}$$

In a third step, the precision of the estimation is examined. Based on the analysis of the variance of the different proportions of publication sets of OpenAPC universities, a confidence interval with an upper and lower 95%-probability threshold is calculated.

3.2 Estimation 2: identification of likely APC-liable publications

The second estimation is more complex as it is based on the identification of publications where APC are likely to be paid for and it considers journal-specific cost information known from other payments included in OpenAPC. In addition, the approach does not take the completeness of OpenAPC for granted but aims to identify additional corresponding author publications of OpenAPC universities for a more complete estimation of costs.

In a first step, a table is created with all journals covered in the WoS for which OpenAPC reports at least one payment. In the case of journals with more than one payment recorded in the period 2018-2019 in OpenAPC, the average APC costs are calculated. In the case of journals without any APC payments in 2018-2019, the most recent payment is selected as an estimation for APC costs in that journal.

In a second step, it is estimated for OpenAPC universities how many additional APC payments are likely to have been made that are not included for that university in OpenAPC. Therefore, it is identified how many corresponding author publications covered by the WoS can be found in journals for which one or more payments are recorded in OpenAPC. In the case of these publications, it is assumed that the journal charges APC and APC payments are likely to have been made outside the publication funds. The (average) payment in the table created in step one is used as a proxy for missing payment records. Finally, the ratio of recorded and unrecorded payments was calculated.

In a third step, the expenses for non-OpenAPC universities are calculated following the same approach as for the OpenAPC universities. For all publications published by a corresponding author of a non-OpenAPC university it is assumed that a payment has been made since the journal charges APC. Again, the (average) costs taken from the table are taken as a proxy for the actual (but unrecorded) payment.

4. Results

This section reports the results of the two estimation procedures described in the previous section.

4.1 Result 1: Global estimation

In a first step, the two alternative versions of the estimation of APC-liable publications that is based on the proportion of different sets of publications are performed. Table 1 refers to OpenAPC universities only and contains all necessary information that is needed for the two versions:

- the number of publications covered by WoS (A),
- the number of publications included in OpenAPC (B),
- publications covered both in WoS and OpenAPC ($A \cap B$)
- the ratio of $A \cap B$ and A,
- the number of publications included both in WoS and in journals of the ISSN-Gold-OA list ($A \cap B$),
- the ratio of $A \cap B$ and A

- the number of publications covered in WoS but not covered in OpenAPC ($B \setminus A$),
- the ratio of publications in OpenAPC not covered in WoS of all publications in OpenAPC ($B \setminus A$ of B), and
- the ratio of all publications in WoS and OpenAPC of publications in WoS and journals of the ISSN-Gold-OA list.

Table 1: OpenAPC universities, intersecting sets and ratio of subsets (2019)

University	A Publ. in WoS	B Publ. in OAPC	$A \cap B$ Publ. OAPC in WoS	$A \cap B$ of A (%)	$A \cap C$ Publ. Gold OA in WoS	$A \cap C$ of A (%)	$B \setminus A$ Publ. OAPC not in WoS	$B \setminus A$ of B (%)	$A \cap B$ of $A \cap C$ (%)
TU München	3,442	460	398	11.56	723	21.01	62	13.48	55.05
LMU München	3,195	89	85	2.66	669	20.94	4	4.49	12.71
Universität Heidelberg	3,009	297	265	8.81	802	26.65	32	10.77	33.04
Erlangen-Nürnberg	2,141	222	193	9.01	453	21.16	29	13.06	42.60
TU Dresden	2,039	265	227	11.13	474	23.25	38	14.34	47.89
KIT Karlsruhe	1,947	232	208	10.68	345	17.72	24	10.34	60.29
Universität Tübingen	1,874	291	262	13.98	458	24.44	29	9.97	57.21
Universität Göttingen	1,774	347	313	17.64	462	26.04	34	9.80	67.75
WWU Münster	1,612	104	92	5.71	321	19.91	12	11.54	28.66
FU Berlin	1,596	112	91	5.70	451	28.26	21	18.75	20.18
Universität Leipzig	1,533	218	187	12.20	381	24.85	31	14.22	49.08
Universität Bochum	1,450	113	104	7.17	289	19.93	9	7.96	35.99
Universität Mainz	1,445	95	82	5.67	333	23.04	13	13.68	24.62
Duisburg-Essen	1,344	168	138	10.27	297	22.10	30	17.86	46.46
TU Berlin	1,118	86	64	5.72	166	14.85	22	25.58	38.55
Universität Stuttgart	1,044	51	43	4.12	133	12.74	8	15.69	32.33
Universität Gießen	1,003	75	69	6.88	269	26.82	6	8.00	25.65
TU Darmstadt	959	73	62	6.47	145	15.12	11	15.07	42.76
Regensburg	948	134	119	12.55	207	21.84	15	11.19	57.49
Universität Rostock	847	99	90	10.63	213	25.15	9	9.09	42.25
Universität Hannover	816	91	78	9.56	157	19.24	13	14.29	49.68
Universität Bremen	788	148	123	15.61	181	22.97	25	16.89	67.96
Universität Potsdam	753	111	96	12.75	183	24.30	15	13.51	52.46
Halle-Wittenberg	721	73	68	9.43	166	23.02	5	6.85	40.96
TU Braunschweig	697	101	83	11.91	155	22.24	18	17.82	53.55
TU Dortmund	633	38	26	4.11	87	13.74	12	31.58	29.89
Universität Bielefeld	569	122	106	18.63	137	24.08	16	13.11	77.37
Universität Konstanz	562	62	56	9.96	124	22.06	6	9.68	45.16
Universität Bayreuth	516	68	58	11.24	102	19.77	10	14.71	56.86
Universität Oldenburg	460	101	85	18.48	132	28.70	16	15.84	64.39
Universität Kassel	427	62	56	13.11	74	17.33	6	9.68	75.68
TU Chemnitz	353	29	26	7.37	64	18.13	3	10.34	40.63
Universität Siegen	295	9	8	2.71	33	11.19	1	11.11	24.24
TU Ilmenau	268	21	15	5.60	37	13.81	6	28.57	40.54
TiHo Hannover	260	108	102	39.23	112	43.08	6	5.56	91.07
Universität Osnabrück	245	37	30	12.24	63	25.71	7	18.92	47.62
Universität Mannheim	241	9	8	3.32	23	9.54	1	11.11	34.78
TU Hamburg-Harburg	237	22	16	6.75	34	14.35	6	27.27	47.06
Universität Bamberg	154	19	12	7.79	14	9.09	7	36.84	85.71

TU Clausthal	128	13	9	7.03	24	18.75	4	30.77	37.50
Universität Passau	94	1	1	1.06	10	10.64	0	0.00	10.00
Total	43,537	4,776	4,154	9.54	9,503	21.83	622	13.02	43.71

For the two alternative estimations, three ratios of sets of publications are used. $A \cap B$ of A , $B \setminus A$ of $A \cap B$ and $A \cap B$ of $A \cap C$. Given that the precision of the estimation depends on the variation of the ratios within the group of OpenAPC universities, confidence intervals are calculated.

Table 2: Variation of proportions

Variable	Observations	Mean	Std. Error	CI (95%) min.	CI (95%) max.
$A \cap B$ of A	41	0.0991	0.0099	0.07913	0.11921
$B \setminus A$ of $A \cap B$	41	0.1462	0.0120	0.1219	0.1704
$A \cap B$ of $A \cap C$	41	0.4619	0.0285	0.4042	0.5196

In the next step, the two ratios (a) of publications covered by OpenAPC and WoS ($A \cap B$) of all publications in WoS (A), and (b) the ratio of publications covered by OpenAPC and not by WoS $B \setminus A$ of all publications covered by OpenAPC (B) are used for an estimation of the number of APC-liable publications of non-OpenAPC universities. The 95%-threshold of the confidence interval was used to calculate likely minimum and maximum numbers of APC-liable publications. In addition, the average APC cost for publications with a German corresponding author taken from OpenAPC was used to estimate total APC costs for each university. This was 1,533 € for 2019.⁷ Again, likely minimum and maximum levels of APC costs were calculated based on the upper and lower limit of the confidence interval of B .

Table 3: Non-OpenAPC universities, observed and estimated values (1)

University	A Publ. in WoS	B est. APC Publ	B _(min) est. APC Publ.	B _(max) est. APC Pubs.	APC est. (€)	APC _(min) est. (€)	APC _(min) est. (€)
RWTH Aachen	2,479	282	220	346	431,837	337,375	529,986
Universität Hamburg	2,193	249	195	306	382,016	298,452	468,842
Universität zu Köln	1,723	196	153	240	300,143	234,489	368,360
Universität Bonn	1,613	183	143	225	280,981	219,518	344,844
Universität Jena	1,286	146	114	179	224,019	175,016	274,934
Universität Düsseldorf	1,173	133	104	164	204,334	159,637	250,776
Universität zu Kiel	1,099	125	98	153	191,444	149,566	234,955
MHH Hannover	875	99	78	122	152,423	119,082	187,066
Universität Magdeburg	660	75	59	92	114,971	89,822	141,102
Universität zu Lübeck	497	56	44	69	86,576	67,638	106,254
Universität Hohenheim	467	53	41	65	81,350	63,556	99,840
TU Kaiserslautern	411	47	36	57	71,595	55,934	87,868
UK Schleswig-Holstein	386	44	34	54	67,240	52,532	82,523
Universität Paderborn	343	39	30	48	59,750	46,680	73,330

⁷ https://treemaps.intact-project.org/apcdata/openapc/#institution/country=DEU&is_hybrid=FALSE&period=2019, retrieved on May 5th 2021.

TU Bergakademie Freiberg	295	34	26	41	51,388	40,147	63,068
Universität Wuppertal	269	31	24	38	46,859	36,609	57,510
Universität Witten/Herdecke	268	30	24	37	46,685	36,473	57,296
Universität Augsburg	250	28	22	35	43,550	34,023	53,448
UK Gießen und Marburg	225	26	20	31	39,195	30,621	48,103
TU Cottbus-Senftenberg	187	21	17	26	32,575	25,449	39,979
Universität Koblenz-Landau	157	18	14	22	27,349	21,367	33,565
Universität Lüneburg	154	17	14	21	26,826	20,958	32,924
Sporthochschule Köln	150	17	13	21	26,130	20,414	32,069
Universität der BW München	139	16	12	19	24,214	18,917	29,717
Jacobs University Bremen	133	15	12	19	23,168	18,100	28,434
FernUniversität in Hagen	94	11	8	13	16,375	12,793	20,096
Universität der BW Hamburg	92	10	8	13	16,026	12,521	19,669
Universität Weimar	85	10	8	12	14,807	11,568	18,172
Universität Eichstätt-Ingolstadt	67	8	6	9	11,671	9,118	14,324
Universität Erfurt	66	7	6	9	11,497	8,982	14,110
Herzzentrum Freiburg	58	7	5	8	10,103	7,893	12,400
Universität Hildesheim	50	6	4	7	8,710	6,805	10,690
Universität Frankfurt (Oder)	35	4	3	5	6,097	4,763	7,483
Universität Vechta	33	4	3	5	5,749	4,491	7,055
Frankfurt School Fin. & Mana.	29	3	3	4	5,052	3,947	6,200
Hertie School of Governance	29	3	3	4	5,052	3,947	6,200
Otto Beisheim School of Mana.	29	3	3	4	5,052	3,947	6,200
MH Brandenburg	27	3	2	4	4,703	3,675	5,772
PH Freiburg	20	2	2	3	3,484	2,722	4,276
Hochschule Musik ... Hannover	16	2	1	2	2,787	2,177	3,421
ESCP Berlin	13	1	1	2	2,265	1,769	2,779
Zeppelin Universität	13	1	1	2	2,265	1,769	2,779
Universität der Künste Berlin	13	1	1	2	2,265	1,769	2,779
PH Ludwigsburg	11	1	1	2	1,916	1,497	2,352
HafenCity Universität Hamburg	11	1	1	2	1,916	1,497	2,352
Psych. HS Berlin	10	1	1	1	1,742	1,361	2,138
PH Heidelberg	10	1	1	1	1,742	1,361	2,138
Universität f. Verwalt. Speyer	7	1	1	1	1,219	953	1,497
PH Karlsruhe	7	1	1	1	1,219	953	1,497
Int. Psych. University Berlin	7	1	1	1	1,219	953	1,497
EBS Wirtschaft und Recht	6	1	1	1	1,045	817	1,283
PH Schwäbisch Gmünd	6	1	1	1	1,045	817	1,283
HS Neuendettelsau	4	0	0	1	697	544	855
Hochschule Hanns Eisler Berlin	4	0	0	1	697	544	855
Hochschule f. Musik Freiburg	4	0	0	1	697	544	855
PH Weingarten	2	0	0	0	348	272	428
Comprehensive Cancer Center	2	0	0	0	348	272	428
Universität Flensburg	1	0	0	0	174	136	214
HS Musik ... München	1	0	0	0	174	136	214
KHS Medien Köln	1	0	0	0	174	136	214
Theologische Fak. Paderborn	1	0	0	0	174	136	214
Steinbeis-Hochschule Berlin	1	0	0	0	174	136	214

Without going too much into the discussion, one can already see from table 3 that the 95%-confidence interval for the estimation of the number of APC-liable publications of non-

OpenAPC universities is broad. This is a result of a large variation of the proportion of the publication sets $A \cap B$ of A within the group of OpenAPC universities.

Therefore, an alternative estimation is performed which estimated APC-liable publications within WoS based on the ratio of publications covered by OpenAPC and WoS ($A \cap B$) and publications in journals that are included in the ISSN-Gold-OA list ($A \cap B$). Similar to the previous method, the ratio of publications covered by OpenAPC and not by WoS (B / A) of all publications covered by OpenAPC (B) are used for an estimation of the number of APC-liable publications of non-OpenAPC universities. Again, the 95%-threshold of the confidence interval was used to calculate likely minimum and maximum numbers of APC-liable publications, and average APC costs were also calculated for each university of that group.

Table 4: Non-OpenAPC universities, observed and estimated values (2)

University	A Publ. in WoS	$A \cap C$	B est. APC Publ	$B_{(min)}$ est. APC Publ.	$B_{(max)}$ est. APC Pubs.	APC est. (€)	APC _(min) est. (€)	APC _(max) est. (€)
RWTH Aachen	2,479	496	263	228	297	402,532	350,156	454,908
Universität Hamburg	2,193	509	269	234	305	413,082	359,333	466,831
Universität zu Köln	1,723	274	145	126	164	222,366	193,433	251,300
Universität Bonn	1,613	345	183	159	206	279,987	243,556	316,418
Universität Jena	1,286	258	137	119	154	209,382	182,137	236,626
Universität Düsseldorf	1,173	279	148	128	167	226,424	196,962	255,886
Universität zu Kiel	1,099	246	130	113	147	199,643	173,666	225,620
MHH Hannover	875	284	150	131	170	230,482	200,492	260,472
Universität Magdeburg	660	163	86	75	98	132,284	115,071	149,496
Universität zu Lübeck	497	117	62	54	70	94,952	82,597	107,307
Universität Hohenheim	467	142	75	65	85	115,241	100,246	130,236
TU Kaiserslautern	411	68	36	31	41	55,186	48,005	62,366
UK Schleswig-Holstein	386	86	46	40	51	69,794	60,712	78,875
Universität Paderborn	343	30	16	14	18	24,347	21,179	27,515
TU Bergakademie Freiberg	295	31	16	14	19	25,158	21,885	28,432
Universität Wuppertal	269	43	23	20	26	34,897	30,356	39,438
Universität Witten/Herdecke	268	70	37	32	42	56,809	49,417	64,201
Universität Augsburg	250	38	20	17	23	30,839	26,826	34,852
UK Gießen und Marburg	225	52	28	24	31	42,201	36,710	47,692
TU Cottbus-Senftenberg	187	26	14	12	16	21,100	18,355	23,846
Universität Koblenz-Landau	157	21	11	10	13	17,043	14,825	19,260
Universität Lüneburg	154	23	12	11	14	18,666	16,237	21,095
Sporthochschule Köln	150	41	22	19	25	33,274	28,944	37,603
Universität der BW München	139	15	8	7	9	12,173	10,589	13,757
Jacobs University Bremen	133	28	15	13	17	22,724	19,767	25,680
FernUniversität in Hagen	94	11	6	5	7	8,927	7,766	10,089
Universität der BW Hamburg	92	11	6	5	7	8,927	7,766	10,089
Universität Weimar	85	11	6	5	7	8,927	7,766	10,089
Universität Eichstätt-Ingolstadt	67	11	6	5	7	8,927	7,766	10,089
Universität Erfurt	66	5	3	2	3	4,058	3,530	4,586
Herzzentrum Freiburg	58	12	6	6	7	9,739	8,472	11,006
Universität Hildesheim	50	9	5	4	5	7,304	6,354	8,254
Universität Frankfurt (Oder)	35	0	0	0	0	0	0	0

Universität Vechta	33	4	2	2	2	3,246	2,824	3,669
Frankfurt School Fin. & Mana.	29	0	0	0	0	0	0	0
Hertie School of Governance	29	2	1	1	1	1,623	1,412	1,834
Otto Beisheim School of Mana.	29	2	1	1	1	1,623	1,412	1,834
MH Brandenburg	27	3	2	1	2	2,435	2,118	2,751
PH Freiburg	20	3	2	1	2	2,435	2,118	2,751
Hochschule Musik ... Hannover	16	5	3	2	3	4,058	3,530	4,586
ESCP Berlin	13	1	1	0	1	812	706	917
Zeppelin Universität	13	1	1	0	1	812	706	917
Universität der Künste Berlin	13	2	1	1	1	1,623	1,412	1,834
PH Ludwigsburg	11	0	0	0	0	0	0	0
HafenCity Universität Hamburg	11	1	1	0	1	812	706	917
Psych. Hochschule Berlin	10	3	2	1	2	2,435	2,118	2,751
PH Heidelberg	10	1	1	0	1	812	706	917
Universität Verwaltung. Speyer	7	0	0	0	0	0	0	0
PH Karlsruhe	7	1	1	0	1	812	706	917
Int. Psych. University Berlin	7	0	0	0	0	0	0	0
EBS Wirtschaft und Recht	6	0	0	0	0	0	0	0
PH Schwäbisch Gmünd	6	1	1	0	1	812	706	917
HS Neuendettelsau	4	0	0	0	0	0	0	0
HS Hanns Eisler Berlin	4	0	0	0	0	0	0	0
HS Musik Freiburg	4	0	0	0	0	0	0	0
PH Weingarten	2	0	0	0	0	0	0	0
Comprehensive Cancer Center	2	1	1	0	1	812	706	917
Universität Flensburg	1	0	0	0	0	0	0	0
HS Musik ... München	1	0	0	0	0	0	0	0
KHS für Medien Köln	1	0	0	0	0	0	0	0
Theol. Fakultät Paderborn	1	0	0	0	0	0	0	0
Steinbeis-Hochschule Berlin	1	0	0	0	0	0	0	0
Musikhochschule Lübeck	0	0	0	0	0	0	0	0

4.2 Estimation 2: Identification of likely APC-liable publications

This section presents the results of the second estimation procedure, which it expected to be more precise for at least three reasons. First, it aims to include APC payments of OpenAPC universities that are not captured by OpenAPC and should therefore be more complete. Second, the calculation of the APC-liable part of the publication output of non-OpenAPC universities is not undertaken on the ground of global proportions of different subsets of publications in a different group of universities (OpenAPC) with considerable variation, but on the identification of individual publications where APC liability is likely. Thus, it should be more context-sensitive. Third, it applies journal-specific payment information instead of average costs.

Table 5 refers to the group of OpenAPC universities. Besides the number of publications with APC payments and the sum of the payments captured in OpenAPC for each university, the number of likely APC-liable publications is given. These publications were published in a journal for which other publications with payment information can be found in OpenAPC. These journals are also covered by the ISSN-Gold-OA list 4.0, indicating that they make all their publications open access.

Table 5: OpenAPC universities, publications covered by OpenAPC and likely APC-liable publications in 2019

University	B Publ. in OAPC	B costs (APC observed) (€)	A ∩ D Likely APC liable publ. in WoS	Costs for A ∩ D (€)	Sum costs B and A ∩ D (€)	Costs not covered by OpenAPC (%)
TU München	460	655,713	355	624,054	1,279,767	48.8
Universität Göttingen	347	537,509	169	306,936	844,445	36.4
Universität Heidelberg	297	468,323	517	974,481	1,442,804	67.5
Universität Tübingen	291	469,584	216	414,796	884,380	46.9
TU Dresden	265	272,230	243	453,250	725,481	62.5
KIT	232	329,661	161	213,433	543,094	39.3
Universität Erlangen-Nürnb.	222	337,001	267	428,829	765,830	56.0
Universität Leipzig	218	342,628	206	358,267	700,895	51.1
Universität Duisburg-Essen	168	260,819	158	266,998	527,817	50.6
Universität Bremen	148	237,189	77	103,454	340,644	30.4
Universität Regensburg	134	245,730	94	156,424	402,153	38.9
Universität Bielefeld	122	186,887	32	47,313	234,200	20.2
Universität Bochum	113	187,325	178	267,035	454,360	58.8
FU Berlin	112	157,778	209	354,521	512,299	69.2
Universität Potsdam	111	167,636	82	133,477	301,113	44.3
TiHo Hannover	108	175,247	16	26,848	202,095	13.3
Universität Münster	104	165,475	235	391,003	556,479	70.3
Universität Oldenburg	101	156,532	46	68,422	224,955	30.4
TU Braunschweig	101	121,605	73	99,766	221,370	45.1
Universität Rostock	99	134,823	105	152,671	287,493	53.1
Universität Mainz	95	152,970	249	375,772	528,742	71.1
Universität Hannover	91	138,968	82	111,480	250,448	44.5
LMU München	89	158,864	559	985,012	1,143,876	86.1
TU Berlin	86	123,275	102	147,500	270,774	54.5
Universität Gießen	75	119,171	204	331,799	450,970	73.6
Universität Halle-Wittenberg	73	116,371	94	152,925	269,296	56.8
TU Darmstadt	73	106,245	85	100,957	207,202	48.7
Universität Bayreuth	68	94,062	45	62,404	156,466	39.9
Universität Konstanz	62	101,493	66	99,849	201,342	49.6
Universität Kassel	62	81,087	16	24,557	105,644	23.3
Universität Stuttgart	51	67,423	87	129,042	196,465	65.7
TU Dortmund	38	47,619	57	71,238	118,857	59.9
Universität Osnabrück	37	59,296	31	51,383	110,680	46.4
TU Chemnitz	29	36,794	43	53,903	90,697	59.4
TU Hamburg-Harburg	22	31,469	17	19,247	50,715	38.0
TU Ilmenau	21	29,560	21	18,360	47,920	38.3
Universität Bamberg	19	31,180	3	3,604	34,784	10.4
TU Clausthal	13	17,825	15	15,369	33,194	46.3
Universität Siegen	9	11,298	29	35,076	46,375	75.6
Universität Mannheim	9	15,880	15	12,486	28,366	44.0
Universität Passau	1	829	8	6,732	7,560	89.0
Total	4,776	7,151,375	5,267	8,650,673	15,802,048	54.7

The approach reveals that the payment data from universities reported to OpenAPC are far from being complete. What was expectable is that individual articles might have been paid

from other sources than the publication funds, for example, because of APC payments not meeting the restrictions of the funding criteria,⁸ overall shortage of money in publication funds or easier accessible funds for the coverage of APC. But the volume of likely APC-liable publications not captured by OpenAPC as well as the total costs come as a surprise. More than 50% of the costs are not covered by OpenAPC in the group of OpenAPC universities, and examples of universities with a larger publication output and large parts of the costs not covered by OpenAPC are LMU München, Universität Heidelberg, FU Berlin, Universität Münster, or TU Dresden. Almost complete coverage is rare, examples are Universität Bamberg or TiHo Hannover.

In a final step of the analysis, the number of likely APC-liable publications is identified and the journal-specific costs for APC are calculated for non-OpenAPC universities. Table 6 summarizes the results.

Table 6: Non-OpenAPC universities, likely APC-liable publications

University	A N D Likely APC liable publ. in WoS	A N D costs (€)
Universität Hamburg	443	803,314
RWTH Aachen	422	741,787
Universität Bonn	302	496,702
MHH Hannover (MHH)	263	521,071
Universität zu Köln	248	487,592
Universität Düsseldorf	246	470,823
Universität Jena	232	425,050
Universität zu Kiel	212	369,026
Universität Magdeburg	136	233,137
Universität Hohenheim	132	180,114
Universität zu Lübeck	108	202,482
UK Schleswig-Holstein	81	149,138
TU Kaiserslautern	62	113,612
Universität Witten/Herdecke	61	114,421
UK Gießen und Marburg	44	87,147
Sporthochschule Köln	41	75,520
Universität Wuppertal	31	54,118
TU Bergakademie Freiberg	29	50,023
Universität Augsburg	29	43,573
Jacobs University Bremen	25	37,530
Universität Paderborn	24	40,959
Universität Koblenz-Landau	20	34,901
Universität Lüneburg	19	30,985
TU Cottbus-Senftenberg	15	16,398
Universität Eichstätt -Ingolstadt	11	20,220
Universität Weimar	8	10,378
Universität der BW München	6	12,152
Universität Hildesheim	6	9,388

⁸ In Germany most notably the criteria of the DFG-funding programme 'Open Access Publizieren' with its limit of 2,000€ for eligible APC (Fournier and Weihberg 2013).

Herzzentrum Freiburg	6	10,202
Universität der BW Hamburg	5	7,716
FernUniversität in Hagen	4	5,976
Universität Vechta	4	6,392
HS für Musik ... Hannover	3	5,469
Universität Erfurt	2	4,041
MHS Brandenburg	2	3,048
PH Freiburg	2	3,002
Otto Beisheim School of Mana.	2	1,730
Comprehensive Cancer Center	1	2,092
PH Karlsruhe	1	1,747
PH Schwäbisch Gmünd	1	1,746
ESCP Berlin	1	1,171
PH Heidelberg	1	1,344
Hertie School of Governance	1	3,128
HafenCity Universität Hamburg	1	1,088
Zeppelin Universität	1	2,020
Total	3,294	5,893,470

5. Discussion

The goal of this article is to compare two methods for the estimation of expenditures for APC of universities that do not contribute to the monitoring system OpenAPC. The first one with its two alternatives is based on the ratio of different subsets of the publication output of universities, the second one on the identification of likely APC-liable publications.

The results reported in the previous section reveal that the two alternative versions of the first approach do not result in a meaningful estimation for at least three reasons: first, the upper and lower limits of the 95% confidence interval define a range of likely APC-liable publications for non-OpenAPC universities that is too broad for financial planning and management of an OA transformation at universities. Second, a comparison of the results of the two alternative versions raises doubts regarding the reliability of the estimation: the confidence interval for four of the 20 universities with the largest publication output covered by the Web of Science does not show any overlap. This is the case for Universität Hohenheim, Universität Paderborn, TU Bergakademie Freiberg and TU Cottbus-Senftenberg. Third, and most important, a comparison of the results of the two versions of the first approach with the second one reveal that all 20 universities with the highest number of likely APC-liable publications have considerably higher estimated APC costs in approach two than in the two versions of approach one outside both confidence intervals. The reason for this is the incompleteness of data from OpenAPC universities that applied for the estimation in the first approach. Therefore, the first approach with its two alternative versions should not even be used as a shortcut for a 'rough' or 'quick and dirty' estimation of APC-liable publications for universities not contributing to OpenAPC.

But how about the second approach, does it result in reliable approximation of APC-liable publications? At first glance, the results seem to be more promising but the approach also has some limitations. First, the identification of likely APC-liable publications is based on

actual APC payments of other publications in the same journal. As a journal may have both changed its OA-business model or the amount of APC, the actual number of APC-liable information and the actual costs may differ from the estimation provided by the second approach. Moreover, it is not possible to decide for the likely APC-liable publications whether or not the estimation is higher or lower than the actual payments that have been made by the university. Second, the additional likely APC-liable publications in the case of OpenAPC universities as well as the likely APC-liable publications in the case of non-OpenAPC universities were identified in the section of the publication output covered by WoS only, but not for the publication output not covered by WoS. Therefore, the actual numbers of likely APC-liable publications are higher. One possible way for an estimation would use the ratio of publications covered by OpenAPC and not covered by WoS and publications covered both by OpenAPC and WoS ($B \setminus A$ of $A \cap B$).⁹

6. Conclusion

The transformation towards OA publishing comes along with the hope for more transparency regarding financial flows on the publication market. Data sets like the OpenAPC give evidence that more transparency is possible. However, the current state of reporting of APC payments shows that there is a considerable blind spot in this monitoring system. This paper tried to investigate this blind spot and compared two approaches for an estimation of costs for APCs paid by German universities that are currently not covered by OpenAPC. The first one is based on the ratio of different subsets in the publication output of German universities that contribute to OpenAPC. It turned out that such a shortcut-approach does not yield reliable results because of the variance of the ratios in the OpenAPC university group and because of the incompleteness of data. Therefore, such an approach should not be used, not even for a quick-and-dirty estimation.

The second approach is based on the identification of likely APC-liable publications and also considers missing publications in the group of OpenAPC universities. The results seem to be more reliable but also come with some limitations, most notably, the coverage of the database. Given that the identification of possible APC-liable publications is based on the WoS, possible APC-liable publications in the publication output not covered by the database cannot be identified. The second approach revealed that OpenAPC payment data are incomplete to a large extent for the group of universities that report to OpenAPC. The reasons for this are beyond the scope of the analysis. Nevertheless, it is likely that local conditions and factors within universities, such as funding criteria of publication funds, limited resources or easier accessible funds (than the central publication funds) may play a role here. This incompleteness of data reported to OpenAPC restricts the value of the data monitoring system and illustrates that more transparency can only be achieved if data reported by universities are exhaustive.

⁹ See Appendix A and B for the results of the estimation.

7. References

- Appel, Andre Luiz, and Sarita Albagli. 'The Adoption of Article Processing Charges as a Business Model by Brazilian Open Access Journals'. *Transinformação* 31 (2019): e180045. <https://doi.org/10.1590/2318-0889201931e180045>.
- Asai, Sumiko. 'Determinants of Article Processing Charges for Medical Open Access Journals'. *Journal of Electronic Publishing* 22, no. 1 (2019). <https://doi.org/10.3998/3336451.0022.103>.
- Barbers, Irene, Nadja Kalinna, and Bernhard Mittermaier. 'Data-Driven Transition: Joint Reporting of Subscription Expenditure and Publication Costs'. *Publications* 6, no. 2 (June 2018): 19. <https://doi.org/10.3390/publications6020019>.
- Björk, Bo-Christer. 'The Hybrid Model for Open Access Publication of Scholarly Articles: A Failed Experiment?' *Journal of the American Society for Information Science and Technology* 63, no. 8 (2012): 1496–1504. <https://doi.org/10.1002/asi.22709>.
- Björk, Bo-Christer, and David Solomon. 'Article Processing Charges in OA Journals: Relationship between Price and Quality'. *Scientometrics* 103, no. 2 (1 May 2015): 373–85. <https://doi.org/10.1007/s11192-015-1556-z>.
- Bruns, Andre, Christopher Lenke, Constanze Schmidt, and Niels Christian Taubert. 'ISSN-Matching of Gold OA Journals (ISSN-GOLD-OA) 4.0', 2020. <https://pub.uni-bielefeld.de/record/2944717>.
- Carroll, Michael W. 'Why Full Open Access Matters'. *PLOS Biology* 9, no. 11 (29 November 2011): e1001210. <https://doi.org/10.1371/journal.pbio.1001210>.
- Crawford, Walt. 'GOAJ2: Gold Open Access Journals 2011-2016', 2017. <https://www.lulu.com/shop/walt-crawford/goaj2-gold-open-access-journals-2011-2016/paperback/product-23188803.html>.
- . *The Gold OA Landscape 2011-2014*. Lulu.com, 2015.
- Gargouri, Yassine, Vincent Larivière, Yves Gingras, Les Carr, and Stevan Harnad. 'Green and Gold Open Access Percentages and Growth, by Discipline'. *ArXiv:1206.3664 [Cs]*, 16 June 2012. <http://arxiv.org/abs/1206.3664>.
- Jahn, Najko, and Marco Tullney. 'A Study of Institutional Spending on Open Access Publication Fees in Germany'. *PeerJ* 4 (9 August 2016): e2323. <https://doi.org/10.7717/peerj.2323>.
- Laakso, Mikael, and Bo-Christer Björk. 'Delayed Open Access: An Overlooked High-Impact Category of Openly Available Scientific Literature'. *Journal of the American Society for Information Science and Technology* 64, no. 7 (2013): 1323–29. <https://doi.org/10.1002/asi.22856>.

- . ‘Hybrid Open Access—A Longitudinal Study’. *Journal of Informetrics* 10, no. 4 (1 November 2016): 919–32. <https://doi.org/10.1016/j.joi.2016.08.002>.
- Laakso, Mikael, Patrik Welling, Helena Bukvova, Linus Nyman, Bo-Christer Björk, and Turid Hedlund. ‘The Development of Open Access Journal Publishing from 1993 to 2009’. *PLoS ONE* 6, no. 6 (13 June 2011). <https://doi.org/10.1371/journal.pone.0020961>.
- Lovén, Lisa. ‘Monitoring Open Access Publishing Costs at Stockholm University’. *Insights* 32, no. 1 (17 January 2019): 3. <https://doi.org/10.1629/uksg.451>.
- Martín-Martín, Alberto, Rodrigo Costas, Thed van Leeuwen, and Emilio Delgado López-Cózar. ‘Evidence of Open Access of Scientific Publications in Google Scholar: A Large-Scale Analysis’. *Journal of Informetrics* 12, no. 3 (1 August 2018): 819–41. <https://doi.org/10.1016/j.joi.2018.06.012>.
- Mongeon, Philippe, and Adèle Paul-Hus. ‘The Journal Coverage of Web of Science and Scopus: A Comparative Analysis’. *Scientometrics* 106, no. 1 (1 January 2016): 213–28. <https://doi.org/10.1007/s11192-015-1765-5>.
- Morrison, Heather, Jihane Salhab, Alexis Calvé-Genest, and Tony Horava. ‘Open Access Article Processing Charges: DOAJ Survey May 2014’. *Publications* 3, no. 1 (March 2015): 1–16. <https://doi.org/10.3390/publications3010001>.
- Pieper, Dirk, and Christoph Broschinski. ‘OpenAPC: A Contribution to a Transparent and Reproducible Monitoring of Fee-Based Open Access Publishing across Institutions and Nations’. *Insights* 31, no. 0 (10 October 2018): 39. <https://doi.org/10.1629/uksg.439>.
- . ‘OpenAPC: A Contribution to a Transparent and Reproducible Monitoring of Fee-Based Open Access Publishing across Institutions and Nations’. *Insights* 31, no. 0 (10 October 2018): 39. <https://doi.org/10.1629/uksg.439>.
- Pinfield, Stephen, Jennifer Salter, and Peter A. Bath. ‘The “Total Cost of Publication” in a Hybrid Open-Access Environment: Institutional Approaches to Funding Journal Article-Processing Charges in Combination with Subscriptions’. *Journal of the Association for Information Science and Technology* 67, no. 7 (2016): 1751–66. <https://doi.org/10.1002/asi.23446>.
- . ‘The “Total Cost of Publication” in a Hybrid Open-Access Environment: Institutional Approaches to Funding Journal Article-Processing Charges in Combination with Subscriptions’. *Journal of the Association for Information Science and Technology* 67, no. 7 (2016): 1751–66. <https://doi.org/10.1002/asi.23446>.
- Piwovar, Heather, Jason Priem, Vincent Larivière, Juan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein. ‘The State of OA: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles’. *PeerJ* 6 (13 February 2018): e4375. <https://doi.org/10.7717/peerj.4375>.
- Prosser, David C. ‘From Here to There: A Proposed Mechanism for Transforming Journals from Closed to Open Access’. *Learned Publishing* 16, no. 3 (2003): 163–66. <https://doi.org/10.1087/095315103322110923>.

- Rimmert, Christine, Holger Schwechheimer, and Matthias Winterhager. 'Disambiguation of Author Addresses in Bibliometric Databases - Technical Report.' Report, 2017. <https://pub.uni-bielefeld.de/record/2914944>.
- . 'Disambiguation of Author Addresses in Bibliometric Databases - Technical Report.' Report, 2017. <https://pub.uni-bielefeld.de/record/2914944>.
- Schönfelder, Nina. 'Article Processing Charges: Mirroring the Citation Impact or Legacy of the Subscription-Based Model?' *Quantitative Science Studies* 1, no. 1 (February 2020): 6–27. https://doi.org/10.1162/qss_a_00015.
- Solomon, David, and Bo-Christer Björk. 'Article Processing Charges for Open Access Publication—the Situation for Research Intensive Universities in the USA and Canada'. *PeerJ* 4 (21 July 2016): e2264. <https://doi.org/10.7717/peerj.2264>.
- Solomon, David J., and Bo-Christer Björk. 'A Study of Open Access Journals Using Article Processing Charges'. *Journal of the American Society for Information Science and Technology* 63, no. 8 (2012): 1485–95. <https://doi.org/10.1002/asi.22673>.
- Suber, Peter. *Open Access*. The MIT Press, 2012. <http://library.oapen.org/handle/20.500.12657/26065>.
- Taubert, Niels, Anne Hobert, Nicolas Fraser, Najko Jahn, and Elham Iravani. 'Open Access -- Towards a Non-Normative and Systematic Understanding'. *ArXiv:1910.11568 [Cs]*, 25 October 2019. <http://arxiv.org/abs/1910.11568>.
- Willinsky, J. 'The Nine Flavours of Open Access Scholarly Publishing'. *Journal of Postgraduate Medicine* 49, no. 3 (7 January 2003): 263.
- Wilson, T.D. 'Re: Bundesrat Decision [Msg. 1078]', 19 April 2007. <http://threader.ecs.soton.ac.uk/lists/boaiforum>.
- Winterhager, Matthias, Holger Schwechheimer, and Christine Rimmert. 'Institutionenkodierung als Grundlage für bibliometrische Indikatoren'. *Bibliometrie - Praxis und Forschung* 3, no. 14 (2014). <https://pub.uni-bielefeld.de/record/2703536>.
- Wohlgemuth, Michael, Christine Rimmert, and Niels Christian Taubert. 'Publikationen in Gold-Open-Access-Journalen auf globaler und europäischer Ebene sowie in Forschungsorganisationen'. Report, 2017. <https://pub.uni-bielefeld.de/record/2912807>.

Appendix A: OpenAPC universities, APC-liable and likely APC-liable publications in WoS with estimated costs and maximum and minimum estimated publications and costs outside (including publications not covered by WoS)

University	B publ in OAPC	B costs (APC observed) (€)	D likely APC liable publ. (all, est. min.)	D likely APC liable publ. (all, est. max.)	D costs (all, est. min.)	D costs (all, est. max.)
TU München	460	655,713	398	415	1,346,020	1,372,380
Universität Göttingen	347	537,509	190	198	875,985	888,534
Universität Heidelberg	297	468,323	580	605	1,539,291	1,577,680
Universität Tübingen	291	469,584	242	253	924,692	940,731
TU Dresden	265	272,230	273	284	770,831	788,875
KIT	232	329,661	181	188	573,141	585,096
Universität Erlangen-Nürnb.	222	337,001	300	312	815,660	835,486
Universität Leipzig	218	342,628	231	241	739,341	754,637
Universität Duisburg-Essen	168	260,819	177	185	557,304	569,036
Universität Bremen	148	237,189	86	90	355,014	360,732
Universität Regensburg	134	245,730	105	110	419,697	426,676
Universität Bielefeld	122	186,887	36	37	240,172	242,549
Universität Bochum	113	187,325	200	208	487,580	500,797
FU Berlin	112	157,778	234	245	551,305	566,824
Universität Potsdam	111	167,636	92	96	316,416	322,505
TiHo Hannover	108	175,247	18	19	205,081	206,269
Universität Münster	104	165,475	264	275	600,336	617,786
Universität Oldenburg	101	156,532	52	54	233,540	236,955
TU Braunschweig	101	121,605	82	85	234,994	240,415
Universität Rostock	99	134,823	118	123	307,090	314,886
Universität Mainz	95	152,970	279	291	575,213	593,702
Universität Hannover	91	138,968	92	96	265,751	271,840
LMU München	89	158,864	627	654	1,248,202	1,289,709
TU Berlin	86	123,275	114	119	289,811	297,384
Universität Gießen	75	119,171	229	239	489,042	504,190
Universität Halle-Wittenberg	73	116,371	105	110	286,839	293,819
TU Darmstadt	73	106,245	95	99	223,066	229,377
Universität Bayreuth	68	94,062	50	53	164,864	168,206
Universität Konstanz	62	101,493	74	77	213,659	218,560
Universität Kassel	62	81,087	18	19	108,630	109,818
Universität Stuttgart	51	67,423	98	102	212,702	219,162
TU Dortmund	38	47,619	64	67	129,495	133,727
Universität Osnabrück	37	59,296	35	36	116,465	118,767
TU Chemnitz	29	36,794	48	50	98,722	101,915
TU Hamburg-Harburg	22	31,469	19	20	53,888	55,150
TU Ilmenau	21	29,560	24	25	51,839	53,398
Universität Bamberg	19	31,180	3	4	35,344	35,567
TU Clausthal	13	17,825	17	18	35,994	37,108
Universität Siegen	9	11,298	33	34	51,787	53,940
Universität Mannheim	9	15,880	17	18	31,166	32,280
Universität Passau	1	829	9	9	9,053	9,647
Total	4,776	7,151,375	5,909	6,164	16,785,022	17,176,116

Appendix B: Non-OpenAPC universities, likely APC-liable publications in WoS and costs with maximum and minimum estimated publications and costs outside WoS

University	A ∩ D likely APC liable publ.	D likely APC liable pub all (est. min.)	D likely APC liable pub all (est. max.)	A ∩ D costs (€)	D costs (all, est. min., €)	D costs (all, est. max. €)
Universität Hamburg	443	497	518	803,314	901,238	940,199
RWTH Aachen	422	473	494	741,787	832,211	868,188
Universität Bonn	302	339	353	496,702	557,249	581,339
MHH Hannover	263	295	308	521,071	584,589	609,861
Universität zu Köln	248	278	290	487,592	547,030	570,678
Universität Düsseldorf	246	276	288	470,823	528,216	551,051
FUniversität Jena	232	260	272	425,050	476,863	497,478
Universität zu Kiel	212	238	248	369,026	414,010	431,908
Universität Magdeburg	136	153	159	233,137	261,556	272,864
Universität Hohenheim	132	148	154	180,114	202,070	210,805
Universität zu Lübeck	108	121	126	202,482	227,165	236,985
UK Schleswig-Holstein	81	91	95	149,138	167,318	174,551
TU Kaiserslautern	62	70	73	113,612	127,461	132,972
Universität Witten/Herdecke	61	68	71	114,421	128,369	133,919
UK Gießen und Marburg	44	49	51	87,147	97,770	101,996
Sporthochschule Köln	41	46	48	75,520	84,726	88,389
Universität Wuppertal	31	35	36	54,118	60,715	63,340
TU Bergakademie Freiberg	29	33	34	50,023	56,121	58,547
Universität Augsburg	29	33	34	43,573	48,884	50,998
Jacobs University Bremen	25	28	29	37,530	42,105	43,925
Universität Paderborn	24	27	28	40,959	45,952	47,938
Universität Koblenz-Landau	20	22	23	34,901	39,156	40,849
Universität Lüneburg	19	21	22	30,985	34,762	36,265
TU Cottbus-Senftenberg	15	17	18	16,398	18,397	19,192
Universität Eichstätt-Ingolstadt	11	12	13	20,220	22,685	23,666
Universität Weimar	8	9	9	10,378	11,643	12,146
Universität der BW München	6	7	7	12,152	13,633	14,223
Universität Hildesheim	6	7	7	9,388	10,533	10,988
Herzzentrum Freiburg	6	7	7	10,202	11,446	11,940
Universität der BW Hamburg	5	6	6	7,716	8,657	9,031
FernUniversität in Hagen	4	4	5	5,976	6,704	6,994
Universität Vechta	4	4	5	6,392	7,171	7,481
Hochschule Musik ... Hannover	3	3	4	5,469	6,135	6,401
Universität Erfurt	2	2	2	4,041	4,533	4,729
MHS Brandenburg	2	2	2	3,048	3,420	3,567
PH Freiburg	2	2	2	3,002	3,368	3,513
Otto Beisheim School of Mana.	2	2	2	1,730	1,940	2,024
Comprehensive Cancer Center	1	1	1	2,092	2,347	2,448
PH Karlsruhe	1	1	1	1,747	1,960	2,044
PH Schwäbisch Gmünd	1	1	1	1,746	1,958	2,043
ESCP Berlin	1	1	1	1,171	1,313	1,370
PH Heidelberg	1	1	1	1,344	1,507	1,573
Hertie School of Governance	1	1	1	3,128	3,510	3,661
HafenCity Universität Hamburg	1	1	1	1,088	1,221	1,274
Zeppelin Universität	1	1	1	2,020	2,267	2,365
Total	3,294	3,696	3,855	5,893,470	6,611,884	6,897,717
