RESEARCH ARTICLE

Who's writing open access (OA) articles? Characteristics of OA authors at Ph.D.-granting institutions in the United States

Anthony J. Olejniczak and Molly J. Wilson

Academic Analytics Research Center (AARC), 1985 W Henderson Road #2159, Columbus, OH 43220

Keywords: article processing charge, bibliometrics, open access, open access publishing, open science, scholarly publishing

ABSTRACT

The open access (OA) publication movement aims to present research literature to the public at no cost and with no restrictions. While the democratization of access to scholarly literature is a primary focus of the movement, it remains unclear whether OA has uniformly democratized the corpus of freely available research, or whether authors who choose to publish in OA venues represent a particular subset of scholars—those with access to resources enabling them to afford article processing charges (APCs). We investigated the number of OA articles with article processing charges (APC OA) authored by 182,320 scholars with known demographic and institutional characteristics at American research universities across 11 broad fields of study. The results show, in general, that the likelihood for a scholar to author an APC OA article increases with male gender, employment at a prestigious institution (AAU member universities), association with a STEM discipline, greater federal research funding, and more advanced career stage (i.e., higher professorial rank). Participation in APC OA publishing appears to be skewed toward scholars with greater access to resources and job security.

1. INTRODUCTION

1.1. Research Objective

Open access (OA) publications present research literature to the public at no cost and with no restrictions. The themes of the open access movement center on research integrity, transparency, and accessibility. Indeed, research funding agencies are more frequently promoting or mandating publication in OA venues and dissemination of code, data, and methods in open repositories (see, e.g., http://roarmap.eprints.org). A fundamental goal of OA is that any person can read published scholarly research, regardless of their ability to pay for access either personally or through an institutional credential. While the democratization of access is a primary driver behind the OA movement (Swan & Brown, 2004; Tennant, Waldner, et al., 2016) it remains unclear whether OA has uniformly democratized the research corpus, or whether a particular subset of authors is more likely to publish their work as OA (specifically, those with access to resources to pay for article processing charges [APCs]).

OA can take many forms, with the most common referred to as a series of colors: Bronze (the article is free to read on the publisher's website but no explicit license is presented); Green (the

an open access journal

Check for updates

Citation: Olejniczak, A. J., & Wilson, M. J. (2020). Who's writing open access (OA) articles? Characteristics of OA authors at Ph.D.-granting institutions in the United States. *Quantitative Science Studies*, *1*(4), 1429–1450. https://doi.org/10.1162/qss_a_00091

DOI: https://doi.org/10.1162/qss_a_00091

Received: 20 March 2020
Accepted: 8 September 2020

Corresponding Author: Anthony J. Olejniczak
aolejniczak@academicanalytics.com

Handling Editor: Ludo Waltman

The MIT Press

article is available in a repository, self-archived by the author); Gold (all articles in the journal are OA); and Hybrid (individual articles are OA if the authors have paid a publication fee, but other articles in the journal are closed). While Bronze and Green represent the majority of all OA publications (Piwowar, Priem, et al., 2018), Gold and Hybrid are unique in that they reflect an author's deliberate decision to make their article immediately publicly available at the time of publication, often paying an APC to do so. In light of these different OA types, we ask two specific questions: (a) What are the characteristics of authors who intend to publish openly immediately (i.e., who choose to publish OA articles), and (b) Which authors are ultimately represented in the OA literature, regardless of the means or type of OA?

1.2. Literature Review

There have been substantial efforts to understand the extent of OA adoption, but none have allowed for granular analysis at the individual author level. Usually, studies focus on the overall number of OA articles published as a percent of the total scholarly literature (E. Archambault, Amyot, et al., 2014; Björk, Welling, et al., 2010; Laakso, Welling, et al., 2011; Piwowar et al., 2018) or the number of journals indexed by OA directories such as ISSN's ROAD (https://road.issn.org/) or the Directory of Open Access Journals (DOAJ) (e.g., Björk, 2019). Others have examined OA by discipline, finding, for example, greater adoption in biomedical research areas (Piwowar et al., 2018). However, discipline-specific OA adoption is often done by a priori classification of journal titles into fields. A disadvantage of this approach is that multidisciplinary journals (e.g., *Nature*, *Science*, *PLOS ONE*) necessitate a time-consuming procedure to subclassify each individual article within those journals, as Piwowar et al. (2018) performed. Additionally, a priori journal classifications fail to account for author-specific discipline affiliations. For example, a scholar whose research program focuses on geochemistry may publish in geology journals, while their academic appointment is in a chemistry department; tagging this scholar's work as only "geology" fails to capture their chemistry discipline affiliation and loses an important characteristic of the individual author.

One frequently cited concern about OA publication is that the method places the burden of publication cost on the scholar, rather than the traditional subscription model, where costs are typically paid by libraries or other resources within a researcher's employing institution. Solomon and Björk (2012) reviewed APCs in OA journals, finding the average APC among OA journals was just over US$900 at the time of their study; similarly, the mean APC was found to be US$899 by Siler and Frenken (2020). In a particularly sharp early criticism of the OA model, Stevenson (2004) wrote in *Times Higher Education*: "The Public Library of Science and the other open-access publishers were created to serve the interests of an elite well-funded and narrow research community." We sought to explore whether authors at more prestigious or wealthy universities publish more often in OA, using the public/private distinction and AAU membership status[1] as the descriptors of institution type.

Another common reservation is the perception that OA journals are less prestigious or have a less scrutinous peer review process (Agrawal, 2014; Beaubien & Eckard, 2014). Although there is evidence that OA research is actually cited more frequently than research behind a paywall (Piwowar et al., 2018), OA publishing is a relatively new phenomenon compared to the approximately 350-year history of traditional scholarly journals (Mabe, 2003). We hypothesize that

[1] The American Association of Universities (http://www.aau.edu) is an invitation-only group of 63 North American research universities (at the time of our study) and membership is widely regarded as evidence of institutional prestige and also access to research support resources. The AAU has been described as "…perhaps the most elite organization in higher education" (Hine, 2010).

young scholars seeking tenure-track positions or promotions may be particularly influenced by the perception of reduced prestige, opting to publish fewer OA articles until their careers are established (i.e., they achieve a tenured position).

There is also evidence that gender affects publication patterns. For example, differences in publishing patterns between men and women are observed in terms of both overall article output (e.g., Cameron, White, & Gray, 2016; Ceci, Ginther, et al., 2014; Duch, Zeng, et al., 2012; Fox, Burns, et al., 2016) and citation patterns, such as a greater proclivity to self-cite among men (King, Bergstrom, et al., 2017). Mishra, Fegley, et al. (2018) found that self-citation differences were largely due to career attrition disproportionately affecting women rather than gender itself as a factor, which underscores that gendered differences in career trajectory play an important role in observed publication and citation patterns. We therefore propose that proclivity to publish in OA journals may also differ between men and women.

In sum, we hypothesize that individual author characteristics (discipline, career stage, overall publication rate, federal research funding support, institution type, and gender) predict an author's likelihood of authoring OA articles in Gold and Hybrid venues, where the deliberate intent to produce an OA article can be inferred. We further hypothesize that the same set of factors predict an author's overall representation in the OA literature (whether Gold, Hybrid, Green, or Bronze). Specifically, we predict that increased institutional prestige, male gender, more advanced career stage, greater overall publishing activity, affiliation with STEM disciplines, and increased federal research grant support will predict increased levels of all types of OA publishing.

2. DATA SOURCES

We culled the names of faculty members at research universities in the United States from the Academic Analytics commercial database (v. AAD2018-1391; http://www.academicanalytics.com), representing faculty rosters for the Fall 2018–Spring 2019 academic year. The Academic Analytics database contains a comprehensive list of faculty members affiliated with one or more academic departments at 390 Ph.D.-granting American research institutions (Supplementary Table 1; https://osf.io/sb8fq/); each department is manually classified by Academic Analytics into one or more of 11 fields of study (Table 1). Academic Analytics uses manual disambiguation and matching to associate each faculty member with their (co)authored CrossRef-DOI journal articles and each of the federal research grants on which they served as a principal investigator (PI). The Academic Analytics database also contains the year of terminal degree for each faculty member, obtained by manual searching. Gender for each faculty member was inferred using genderize.io with a 95% threshold (https://genderize.io). Genderize.io uses a large, global sample of given names associated with known gender gleaned from public social profiles to assign gender probabilities. We used genderize.io to infer gender for the first and middle name of each scholar in the Academic Analytics database; in cases where the first name resulted in a lower than 95% probability of inferring the gender, we deferred to the result for the middle name. We kept in our sample only faculty members with an inferred gender of "male" or "female" based on first name (or middle name when the first name resulted in "unknown"). We assembled the following attributes for each faculty member:

1. A binary indicator whether the faculty member has won federal research grant funding over a 5-year period (2014–2018).
2. The total number of journal articles authored or coauthored over five years (2014–2018), including the DOI, journal name, ISSN, and EISSN for each article for which the scholar is listed as an author.

Table 1. Scholar and authorship counts by field

Field	Number of unique scholars	Scholars with at least one authorship of any kind	Scholars with at least one APC OA authorship
Agricultural Sciences	4,860	4,510 (92.8%)	3,448 (70.9%)
Biological and Biomedical Sciences	30,725	29,207 (95.1%)	25,930 (84.4%)
Business	11,764	9,911 (84.2%)	1,594 (13.5%)
Education	8,317	6,466 (77.7%)	1,936 (23.3%)
Engineering	22,658	20,763 (91.6%)	14,753 (65.1%)
Family, Consumer and Human Sciences	13,609	10,072 (74.0%)	4,808 (35.3%)
Health Professions Sciences	13,098	11,448 (87.4%)	7,738 (59.1%)
Humanities	34,159	19,852 (58.1%)	3,541 (10.4%)
Natural Resources and Conservation	3,913	3,664 (93.6%)	2,977 (76.1%)
Physical and Mathematical Sciences	30,415	27,843 (91.5%)	19,954 (65.6%)
Social and Behavioral Sciences	28,228	24,848 (88.0%)	10,076 (35.7%)

3. The professorial rank of each scholar (assistant professor, associate professor, full professor; while not always true, an associate or full professor title typically indicates a tenured position, whereas assistant professors are usually not tenured).
4. The year each scholar completed their terminal degree (typically Ph.D., sometimes M.F.A., M.B.A., etc.).
5. An indicator of the scholar's employing institution's status as a public or private university.
6. An indicator of the scholar's employing institution's status as a member of the Association of American Universities (AAU; http://www.aau.edu).
7. The inferred gender of each scholar.

We downloaded the contents of the Unpaywall database in December, 2019 (http://unpaywall.org/; Piwowar et al., 2018). Unpaywall harvests article information from OA repositories and other sources, including CrossRef, PubMed Central, and DOAJ. We matched the DOI of articles matched to scholars in the Academic Analytics database with the DOI from the Unpaywall database to classify each journal article (co-)authored by a researcher in our study sample as either "Closed" (not OA), Bronze, Green, Gold, or Hybrid. Over the 5-year study period, 1,585,176 (98%) of the journal articles matched to scholars in our extract from Academic Analytics were also in Unpaywall with their OA status indicated.

The Gold and Hybrid OA models both imply that the authors made an explicit decision to place their article in an OA journal, agreeing to pay APCs and making their research available immediately upon publishing. We recognize that not all Gold OA journals charge APCs, but the Unpaywall data do not distinguish non-APC from APC Gold journals. Moreover, a majority of Gold OA articles are published in venues that do charge APCs (see discussion by Crotty, 2015). Our research questions focus on (a) an individual author's decision to publish an article immediately upon publication, including their willingness to absorb APCs, which is captured by their combined Gold and Hybrid article count, and (b) an individual author's representation in the OA literature as a whole, through any mechanism or OA color. Moreover, Unpaywall's definitions of OA types distinguish Green

from Hybrid OA explicitly: An article is Green if the publisher's website charges a fee for the article but the article is also freely available in a repository; an article is Hybrid if there is no such fee on the publisher's website to read the article (whether or not the article is also placed in a repository). Thus, we summarized each author's article production as follows: The sum of Gold and Hybrid articles is the author's count of "APC OA articles" and the sum of all Gold, Hybrid, Bronze, and Green articles is an author's count of "Total OA articles."

3. REGRESSION MODEL

Our model is designed to analyze how information describing a scholar (gender, professorial rank, years since terminal degree, federal research funding) and characteristics describing their institutional employer (public or private university, AAU member institution) are associated with the number of OA journal articles those scholars author (both APC OA and Total OA). As OA articles published is a discrete variable and the number of scholars who published zero OA articles in the 5-year study period shows a high level of zero-inflation in each field (Table 1), an ordinary least squares regression is inappropriate. General linear models (GLMs) suitable for discrete dependent variables include the Poisson model and the negative binomial model, but neither is appropriate for data with substantial zero-inflation. A hurdle model (Mullahy, 1986; Zeileis, Kleiber, & Jackman, 2008), however, can account for zero-inflation by framing the model as two components: One component assumes a binomial probability model governs the binary outcome of whether a count measurement (in our case, the count of APC OA or Total OA articles) is either zero or greater than zero, depending on whether a "hurdle" has been overcome; the second component fits a truncated-at-zero count model for the positive OA article count observations (i.e., those cases for which the hurdle to publishing at least one OA article has been overcome).

We propose that zero-inflation in the APC OA or total OA article count data reflects a hurdle not overcome by scholars, related to (a) a lack of publishing activity overall (scholars with zero total articles published over our sample's 5-year period may represent cases where publishing research articles is not incentivized or is not part of their research dissemination strategy) or (b) a lack of federal funding supporting one's research; public research funders increasingly mandate that grant recipients publish in OA venues so federal support may be a strong predictor of nonzero OA article counts. We conducted two hurdle models for each of the 11 fields of study, using the *pscl* package (Jackman, 2017; Zeileis et al., 2008) in *R* version 3.6.1 (R Development Core Team, 2011). The first model predicts APC OA articles authored; the second model predicts total OA articles authored.

In the zero count component of each model, we used "APC OA articles" or "Total OA articles" as the dependent variable (interpreted as a binary variable with value "0" for zero APC OA or Total OA articles and value "1" for greater than zero APC OA or Total OA articles), and the following independent variables:

1. Gender (female = 0, male = 1)
2. Professorial rank (as of the Fall 2018–Spring 2019 academic year (assistant professor = 1, associate professor = 2; full professor = 3)
3. Public or private institution (private = 0, public = 1)
4. AAU member institution (not AAU member = 0, AAU member = 1)
5. The total number of articles (combining OA and non-OA) published in the 5-year study period
6. Years since terminal degree (e.g., if a scholar earned their Ph.D. in 2008, years since terminal degree = 2018 − 2008 = 10 years)

7. A binary "Supported by Federal Funding" indicator (has won a federal research grant in the 5-year period = 1; has not won a federal research grant in the 5-year period = 0); only principle investigators on research grants are indicated, other roles (e.g., key personnel) were not available from our data source

For the positive count component, we again used "APC OA articles" or "Total OA articles" as the dependent variable, and the same set of seven independent variables as the zero count component. To explore dispersion in the regression model, we compared Poisson and negative binomial distributions to the actual OA count frequencies. Following Kleiber and Zeileis (2016), we visualized both observed and expected frequencies of OA article counts as hanging rootograms (Supplementary Figure 1; https://osf.io/sb8fq/). In all fields, a negative binomial distribution was a better fit than a Poisson distribution, and results reported hereafter are from the negative binomial models.

4. RESULTS

4.1. Descriptive Statistics

We dropped 18,164 (9.3%) cases from the original sample of 200,484 scholars for whom gender was inferred as "Unknown." Of the 182,320 unique faculty members in the database with an inferred gender of "female" or "male," 65,294 (35.8%) are inferred to be "female" and 117,026 (64.2%) are inferred to be "male." The data set contains 84,401 full professors (46.3%), 55,499 associate professors (30.4%), and 42,523 assistant professors (23.3%). The sample included 77,905 scholars at AAU member institutions (42.7%) and 104,611 scholars at non-AAU member institutions (57.3%). Private institutions employ 48,214 of the scholars in our sample (26.4%), while 134,237 scholars are at public institutions (73.6%). Table 2 gives summary information for each of the 11 fields. Scholars were each affiliated with one or more of 11,436 university departments or other academic units, and each academic unit was classified into one or more of the 11 fields of study. The complete list of academic units and classifications is given as Supplementary Table 1 (https://osf.io/sb8fq/).

Individuals in the study sample authored 1,618,502 journal articles in 25,894 academic journals between 2014 and 2018 (Supplementary Table 2; https://osf.io/sb8fq). The number of scholars who authored articles in the 5-year period is given for each field in Table 1. Of the articles authored by scholars in our study sample, 868,817 are "Closed" (53.7%), 197,791 are Gold OA (12.2%; all articles in this journal are OA), 253,024 are Green OA (15.6%; the article is available in a repository, self-archived by the author), 175,248 are Bronze OA (10.8%; the article is free to read on the publisher's website but no explicit license is presented), and 123,622 are Hybrid OA (7.6%; individual articles within the journal are OA if the authors have paid an APC). The percentages of authorships by OA classification and by demographic and institutional characteristics of the author are shown in Figure 1. The percentage of each group's authorships that are APC OA ranges from 22% of authorships by women to 29 by scholars at non-AAU member institutions. Scholars at private institutions author 2% more of their articles in an APC OA venue than those at public institutions. Scholars at AAU member institutions author 5% more of their articles in APC OA venues than those at non-AAU member institutions. Men author 6% more of their articles in APC OA venues than women. Full professors author 3% more of their articles in APC OA venues than either associate or assistant professors.

The 11 fields we examined reveal a wide range of adoption of APC OA publishing (Table 1; Figure 2). Among Humanities and Business scholars, only 10.4% and 13.5% have authored an APC OA article in the 5-year study period, respectively. The greatest rate of participation in

Table 2. Scholar counts by gender, professorial rank, and institution type in each field

	Female scholars	Male scholars	Scholars at AAU institutions	Scholars at non-AAU institutions	Scholars at private institutions	Scholars at public institutions	Assistant Professors	Associate Professors	Full Professors
Agricultural Sciences	1,394	3,466	2,083	2,777	146	4,714	1,158	1,173	2,529
Biological and Biomedical Sciences	9,937	20,788	14,217	16,508	10,518	20,207	7,166	8,042	15,517
Business	3,268	8,496	4,755	7,009	3,662	8,102	3,115	3,566	5,083
Education	4,984	3,333	2,482	5,835	1,449	6,868	2,222	3,055	3,040
Engineering	4,008	18,650	10,977	11,681	5,512	17,146	5,240	5,719	11,699
Family, Consumer and Human Sciences	6,467	7,142	5,410	8,199	2,567	11,042	3,722	4,896	4,991
Health Professions Sciences	7,862	5,236	5,366	7,732	3,340	9,758	4,374	4,046	4,678
Humanities	14,812	19,347	14,896	19,263	9,906	24,253	6,223	13,077	14,859
Natural Resources and Conservation	1,072	2,841	1,457	2,456	472	3,441	901	1,066	1,946
Physical and Mathematical Sciences	6,339	24,076	13,758	16,657	7,898	22,517	6,287	7,360	16,768
Social and Behavioral Sciences	11,540	16,688	11,569	16,659	7,922	20,306	6,413	8,770	13,045

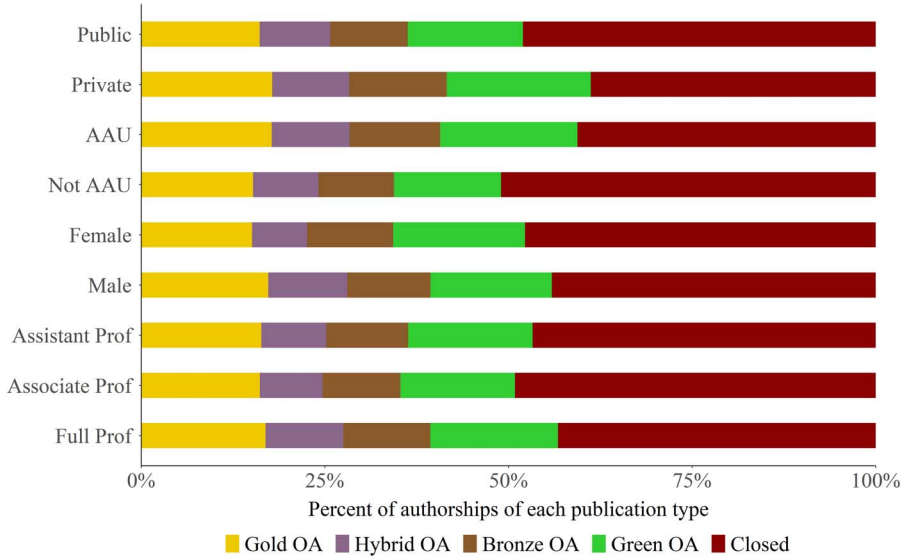

Figure 1. Percentage of articles by OA type by institutional characteristics, gender, and professorial rank.

APC OA publishing is found among scholars in Biological and Biomedical Sciences and Natural Resources and Conservation, among whom 84.4% and 76.1% have authored at least one APC OA article. In five of the 11 fields, fewer than 50% of scholars have authored an APC OA article in the previous 5 years. Figure 2 shows the percentage of authorships in each field by OA type, revealing that even in fields where the rate of participation in APC OA publishing is high, the proportion of articles published that are APC OA (Gold and Hybrid) never reaches 50%. In Physical and Mathematical Sciences, for example, 65.6% of authors have published at least one APC OA article in the previous 5 years, but only 43% of articles authored are APC OA. Likewise, among Agricultural Sciences researchers, 70.0% of scholars have authored an APC OA article, but only 24.7% of their articles are APC OA.

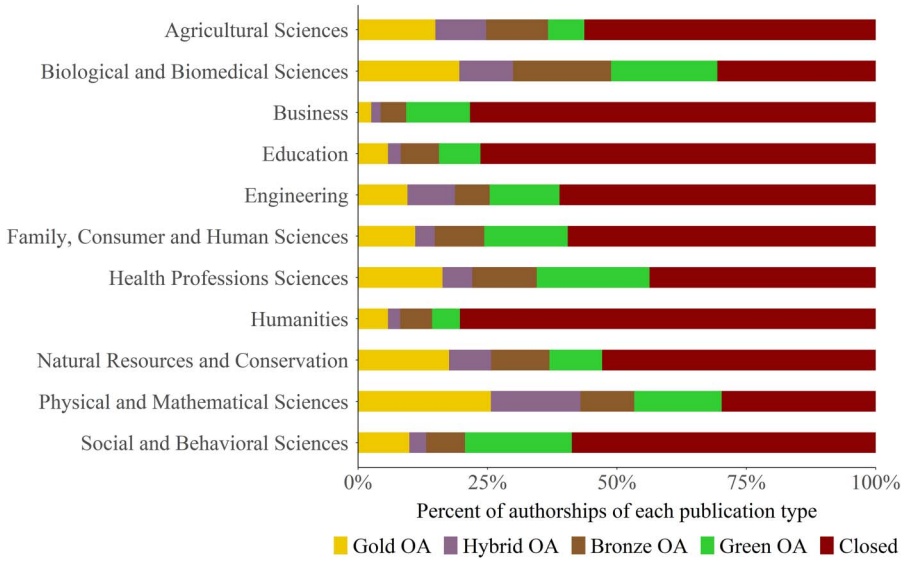

Figure 2. Percentage of articles by OA type published by scholars in each field.

4.2. Regression Results—Predicting APC OA Articles

The results of the APC OA hurdle models by field are given in Tables 3, 4, and 5 for both the truncated positive counts component (predicting the number of APC OA articles authored among scholars who have authored at least one APC OA article) and the zero count component (predicting the binary variable "scholar has/has not authored at least one APC OA article." The exponentiated coefficients for the truncated positive counts component are displayed in Figure 3.

4.2.1. Zero count component of hurdle model: APC OA articles

In all fields, the total number of articles authored was a positive and significant predictor of having authored at least one APC OA article, suggesting that increased overall publication activity increases the likelihood of authoring at least some of those articles in Gold and Hybrid OA venues. For each additional article authored, the likelihood of having authored at least one APC OA article increases by a factor between 1.12 (Business) and 1.29 (Natural Resources and Conservation). Securing federal research grants over the 5-year study period was also a positive and significant predictor of authoring at least one APC OA article in all 11 fields. Most fields show a pronounced increase in the likelihood of authoring APC OA articles if federal research grants were won, including two fields in which the likelihood of publishing at least one APC OA article increases by a factor more than 2.0 if a scholar has won a federal grant (Tables 3–5). AAU membership was a consistently strong predictor of having authored at least one APC OA article; researchers at AAU member institutions are significantly more likely to have published at least one APC OA article than their non-AAU counterparts in five fields. In one field (Engineering), however, AAU membership had a significant negative relationship with having authored at least one APC OA article.

Men are significantly more likely than women to have authored at least one APC OA article in three fields (Biological and Biomedical Sciences, Education, and Health Professions Sciences), and women are significantly more likely than men to have authored at least one APC OA article in three fields (Engineering; Family, Consumer, and Human Sciences; and Natural Resources and Conservation). Scholars at public institutions are significantly less likely to have authored at least one APC OA article than their peers at private institutions in three fields (Agricultural Sciences, Business, and Engineering), but public institution scholars are more likely to have authored at least one APC OA article than private institution scholars in two fields (Health Professions Sciences and Social and Behavioral Sciences). Years since degree has a significant negative relationship with having authored at least one APC OA article in four fields (Biological and Biomedical Sciences, Business, Engineering, Humanities, and Physical and Mathematical Sciences), and a positive relationship with having authored at least one APC OA article in one field (Social and Behavioral Sciences). Increased professorial rank (from assistant professor, to associate professor, to full professor) has a positive significant relationship with publishing at least one APC OA article in one field (Humanities) and a significant negative relationship in three fields (Agricultural Sciences, Engineering, and Social and Behavioral Sciences).

4.2.2. Truncated positive count component of hurdle model: APC OA articles

In every field, federal research grant support and the total count of articles authored were positive significant predictors of the number of APC OA articles authored (Tables 3–5; Figure 3). Having won at least one federal research grant results in an expected increase in the number of APC OA articles authored by a factor of between 1.19 (Health Professions Sciences and Natural Resources and Conservation) to 1.79 (Business). Each additional article authored increases the likelihood of authoring an additional APC OA article by a factor of between 1.02 (Biological and Biomedical Sciences and Physical and Mathematical Sciences) and 1.16 (Humanities). Affiliation with an AAU institution has a positive significant relationship with the number of APC OA articles authored

Table 3. Regression results from the APC OA hurdle models for Agricultural Sciences, Biological and Biomedical Sciences, Business, and Education, showing the count data component (APC OA articles > 0) and the zero count component (APC OA articles = 0)

	Agricultural Sciences				Biological & Biomedical Sci.				Business				Education			
	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.
	Positive Count Component				Positive Count Component				Positive Count Component				Positive Count Component			
Gender is Male	0.094	*	1.099	0.040	0.102	***	1.107	0.011	0.058		1.059	0.654	0.285	***	1.329	0.085
Institution is Public	−0.187	*	0.829	0.091	−0.029		0.971	0.010	−0.090		0.914	0.459	−0.035		0.965	0.109
Institution is AAU Member	−0.006		0.994	0.036	0.095	***	1.100	0.010	−0.022		0.978	0.855	−0.003		0.997	0.089
Years Since Degree	−0.003		0.997	0.002	−0.009	***	0.991	0.001	0.022	***	1.022	0.001	0.003		1.003	0.006
Professorial Rank	−0.005		0.995	0.032	0.147	***	1.158	0.009	−0.327	**	0.721	0.001	−0.040		0.961	0.076
Has Won Research Grants	0.314	***	1.368	0.038	0.223	***	1.249	0.010	0.581	**	1.788	0.002	0.278	*	1.321	0.109
Total Article Count	0.040	***	1.041	0.001	0.024	***	1.024	0.000	0.087	***	1.091	0.000	0.070	***	1.073	0.007
Log(θ)	0.409	***	1.506	0.054	1.000	***	2.718	0.016	−4.122		0.016	0.387	−1.038	***	0.354	0.309
Constant	0.547	***	1.728	0.110	0.788	***	2.198	0.019	−4.554		0.011	4.730	−1.745	***	0.175	0.306
	Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component			
Gender is Male	0.039		1.040	0.086	0.116	**	1.123	0.041	−0.045		0.956	0.067	0.168	**	1.183	0.058
Institution is Public	−0.521	*	0.594	0.264	0.023		1.023	0.043	−0.241	***	0.786	0.064	0.095		1.100	0.077
Institution is AAU Member	0.010		1.010	0.081	0.133	**	1.143	0.041	−0.004		0.996	0.063	0.089		1.093	0.062
Years Since Degree	0.004		1.004	0.005	−0.013	***	0.987	0.002	0.004		1.004	0.003	−0.006		0.994	0.004
Professorial Rank	−0.143	*	0.866	0.073	−0.005		0.995	0.034	0.024		1.024	0.054	−0.016		0.984	0.051
Has Won Research Grants	0.381	***	1.463	0.111	0.723	***	2.060	0.047	1.397	***	4.045	0.134	0.263	**	1.300	0.088
Total Article Count	0.206	***	1.229	0.008	0.250	***	1.284	0.005	0.113	***	1.119	0.004	0.129	***	1.138	0.005
Constant	−0.345		0.708	0.292	−0.482	***	0.618	0.069	−2.680	***	0.069	0.112	−2.062	***	0.127	0.110
Observations	4,860				30,725				11,764				8,317			
Log-likelihood	−9,954				−75,120				−5,786				−6,228			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 4. Regression results from the APC OA hurdle models for Engineering, Family, Consumer, and Human Sciences, Health Professions Sciences, and Humanities, showing the count data component (APC OA articles > 0) and the zero count component (APC OA articles = 0)

	Engineering				Fam., Cons., & Hum. Sci.				Health Professions Sci.				Humanities			
	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.
	Positive Count Component				Positive Count Component				Positive Count Component				Positive Count Component			
Gender is Male	−0.005		0.995	0.021	0.073	*	1.076	0.035	0.277	***	1.320	0.023	0.085		1.088	0.078
Institution is Public	−0.091	***	0.913	0.019	−0.081		0.922	0.046	−0.077	**	0.925	0.026	−0.086		0.918	0.083
Institution is AAU Member	0.086	***	1.090	0.017	0.085	*	1.088	0.036	0.212	***	1.236	0.024	−0.088		0.916	0.081
Years Since Degree	−0.004	***	0.996	0.001	0.003		1.003	0.002	−0.004	**	0.996	0.001	0.011	*	1.012	0.005
Professorial Rank	0.023		1.023	0.015	−0.073	*	0.929	0.032	0.019		1.019	0.020	−0.133		0.875	0.071
Has Won Research Grants	0.237	***	1.267	0.017	0.367	***	1.443	0.040	0.173	***	1.189	0.024	0.372	**	1.451	0.121
Total Article Count	0.029	***	1.030	0.000	0.033	***	1.034	0.001	0.030	***	1.031	0.001	0.148	***	1.160	0.011
Log(θ)	0.546	***	1.727	0.027	0.102		1.108	0.061	0.404	***	1.497	0.035	−2.951	**	0.052	1.010
Constant	0.396	***	1.486	0.036	0.060		1.061	0.074	0.364	***	1.439	0.044	−3.872	***	0.021	1.006
	Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component			
Gender is Male	−0.149	**	0.862	0.046	−0.089		0.915	0.047	0.603	***	1.827	0.050	−0.018		0.982	0.040
Institution is Public	−0.079		0.924	0.042	0.061		1.063	0.062	0.206	***	1.228	0.057	0.002		1.002	0.043
Institution is AAU Member	−0.086	*	0.917	0.036	0.220	***	1.246	0.048	0.336	***	1.399	0.051	0.228	***	1.256	0.040
Years Since Degree	−0.008	***	0.992	0.002	−0.005		0.995	0.003	0.000		1.000	0.003	−0.017	***	0.983	0.002
Professorial Rank	−0.090	**	0.914	0.032	−0.049		0.952	0.041	−0.034		0.967	0.040	0.099	**	1.104	0.037
Has Won Research Grants	0.469	***	1.598	0.039	0.500	***	1.649	0.081	0.395	***	1.484	0.067	0.464	***	1.591	0.077
Total Article Count	0.164	***	1.178	0.003	0.176	***	1.193	0.004	0.179	***	1.195	0.004	0.243	***	1.276	0.005
Constant	−0.715	***	0.489	0.071	−2.007	***	0.134	0.088	−1.880	***	0.153	0.079	−2.957	***	0.052	0.076
Observations	22,658				13,609				13,098				34,159			
Log-likelihood	−42,670				−15,330				−24,050				−12,790			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 5. Regression results from the APC OA hurdle models for Natural Resources and Conservation, Physical and Mathematical Sciences, and Social and Behavioral Sciences, showing the count data component (APC OA articles > 0) and the zero count component (APC OA articles = 0)

	Nat. Res. & Conservation				Phys. & Math. Sci.				Soc. & Behav. Sci.			
	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.
	Positive Count Component				Positive Count Component				Positive Count Component			
Gender is Male	0.031		1.031	0.036	0.069	**	1.072	0.022	0.138	***	1.147	0.032
Institution is Public	−0.035		0.966	0.045	−0.079	***	0.924	0.021	0.043		1.044	0.034
Institution is AAU Member	0.038		1.039	0.032	0.174	***	1.190	0.019	0.145	***	1.156	0.032
Years Since Degree	−0.005	**	0.995	0.002	−0.003	*	0.997	0.001	0.007	***	1.007	0.002
Professorial Rank	0.021		1.021	0.028	0.126	***	1.134	0.016	−0.157	***	0.855	0.028
Has Won Research Grants	0.177	***	1.193	0.032	0.225	***	1.253	0.018	0.374	***	1.453	0.034
Total Article Count	0.036	***	1.036	0.001	0.018	***	1.018	0.000	0.061	***	1.063	0.002
Log(θ)	0.967	***	2.631	0.054	−0.348	***	0.706	0.027	−0.653	***	0.521	0.063
Constant	0.653	***	1.921	0.070	0.368	***	1.446	0.040	−0.919	***	0.399	0.075
	Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component			
Gender is Male	−0.208		0.812	0.109	−0.011		0.989	0.037	0.044		1.045	0.030
Institution is Public	0.167		1.182	0.165	0.050		1.052	0.035	0.099	**	1.104	0.034
Institution is AAU Member	−0.079		0.924	0.106	0.016		1.016	0.032	0.185	***	1.203	0.031
Years Since Degree	−0.002		0.998	0.006	−0.010	***	0.990	0.002	0.003		1.003	0.002
Professorial Rank	−0.129		0.879	0.088	−0.008		0.992	0.027	−0.117	***	0.890	0.028
Has Won Research Grants	0.280	*	1.323	0.123	0.267	***	1.306	0.033	0.686	***	1.985	0.041
Total Article Count	0.254	***	1.289	0.011	0.177	***	1.194	0.003	0.141	***	1.152	0.002
Constant	−0.919	***	0.399	0.224	−0.953	***	0.385	0.060	−1.927	***	0.146	0.056
Observations	3,913				30,415				28,228			
Log-likelihood	−8,287				−65,540				−32,000			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

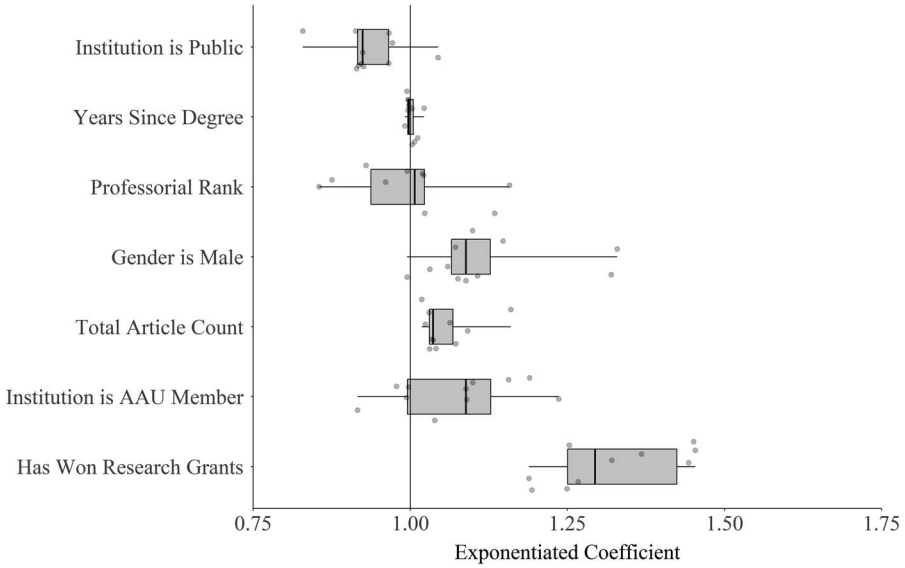

Figure 3. Range of exponentiated coefficients from the positive count component of the APC OA articles hurdle model for each independent variable. Values greater than the vertical reference line at 1.0 represent a greater likelihood of publishing APC OA articles; values less than 1.0 represent a lower likelihood of publishing APC OA articles. Individual dots represent the actual value for each of the 11 fields that constitute the box and whisker summary. The line in the center of each box is the median, the ends of the boxes are the 25th and 75th percentiles, and the whiskers represent the full data range.

in six fields, by a factor ranging between 1.09 (Engineering and Family, Consumer, and Human Sciences) and 1.24 (Health Professions Sciences).

Scholars affiliated with public universities are expected to author fewer APC OA articles than their peers at private universities in 10 of the 11 fields (Tables 3–5), and this relationship is significant in five of those fields. Agricultural Sciences shows the lowest exponentiated coefficient, with public institution scholars authoring APC OA articles at a factor of 0.83 that of their private institution colleagues. In one field (Social and Behavioral Sciences) scholars at public institutions are expected to author slightly more APC OA articles than their peers at private institutions, although this effect is not significant (Table 5). In 10 of the 11 fields, men author more APC OA articles than women, and this result is significant in seven fields (Tables 3–5). The strongest effect is in Education, where men are expected to author more APC OA articles than women by a factor of 1.33 (Table 3). In one field (Engineering), women are expected to author more APC OA articles then men, although this effect is not significant (Table 4).

Years since terminal degree is a positive significant predictor of APC OA articles authored in three fields (Business, Humanities, and Social and Behavioral Sciences), and a negative significant predictor in five fields (Tables 3–5). In four fields, increased professorial rank has a significant negative relationship with APC OA articles authored (Business; Family, Consumer, and Human Sciences; Humanities; and Social and Behavioral Sciences), and in two fields increased rank has a positive relationship with APC OA articles authored (Biological and Biomedical Sciences and Physical and Mathematical Sciences).

4.3. Regression Results—Predicting Total OA Articles

The results of the Total OA articles hurdle models by field are given in Tables 6, 7, and 8 for both the truncated positive counts component (predicting the total number of OA articles published among

Table 6. Regression results from the Total OA Articles hurdle models for Agricultural Sciences, Biological and Biomedical Sciences, Business, and Education, showing the count data component (Total OA articles > 0) and the zero count component (Total OA articles = 0)

	Agricultural Sciences				Biological & Biomedical Sci.				Business				Education			
	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.
	Positive Count Component				Positive Count Component				Positive Count Component				Positive Count Component			
Gender is Male	0.027		1.027	0.028	0.052	***	1.054	0.007	0.056		1.058	0.041	0.129	**	1.138	0.046
Institution is Public	−0.127		0.881	0.068	−0.058	***	0.943	0.007	−0.256	***	0.774	0.038	−0.102		0.903	0.058
Institution is AAU Member	0.077	**	1.081	0.026	0.125	***	1.133	0.007	0.428	***	1.534	0.039	0.065		1.067	0.048
Years Since Degree	−0.002		0.998	0.002	−0.006	***	0.994	0.000	−0.007	**	0.993	0.002	0.003		1.003	0.003
Professorial Rank	−0.003		0.997	0.023	0.115	***	1.121	0.006	0.092	**	1.096	0.033	−0.052		0.949	0.041
Has Won Research Grants	0.323	***	1.381	0.028	0.291	***	1.337	0.007	0.517	***	1.677	0.071	0.209	***	1.233	0.061
Total Article Count	0.041	***	1.042	0.001	0.028	***	1.028	0.000	0.080	***	1.083	0.003	0.089	***	1.093	0.004
Log(θ)	0.876	***	2.400	0.039	1.483	***	4.408	0.012	0.265	***	1.304	0.071	−0.017		0.983	0.093
Constant	1.036	***	2.818	0.081	1.456	***	4.290	0.013	−0.616	***	0.540	0.080	−0.526	***	0.591	0.104
	Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component			
Gender is Male	−0.121		0.886	0.115	−0.073		0.930	0.068	−0.020		0.980	0.049	0.204	***	1.226	0.055
Institution is Public	−0.615		0.541	0.392	−0.019		0.981	0.068	−0.346	***	0.708	0.049	0.031		1.032	0.073
Institution is AAU Member	0.031		1.031	0.109	0.344	***	1.411	0.068	0.470	***	1.600	0.046	0.050		1.052	0.061
Years Since Degree	0.005		1.005	0.007	−0.006		0.994	0.003	−0.008	**	0.992	0.003	0.005		1.005	0.003
Professorial Rank	−0.113		0.893	0.095	−0.059		0.942	0.054	0.127	**	1.135	0.041	−0.055		0.946	0.048
Has Won Research Grants	0.270		1.310	0.171	0.835	***	2.306	0.100	0.985	***	2.678	0.179	0.154		1.167	0.096
Total Article Count	0.464	***	1.590	0.019	0.858	***	2.360	0.019	0.249	***	1.282	0.006	0.299	***	1.349	0.008
Constant	−0.339		0.713	0.426	−0.964	***	0.382	0.112	−1.652	***	0.192	0.083	−1.784	***	0.168	0.104
Observations	4,860				30,725				11,764				8,317			
Log-likelihood	−11,950				−92,020				−14,420				−9,984			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 7. Regression results from the Total OA Articles hurdle models for Engineering, Family, Consumer, and Human Sciences, Health Professions Sciences, and Humanities, showing the count data component (Total OA articles > 0) and the zero count component (Total OA articles = 0)

	Engineering				Fam., Cons., & Hum. Sci.				Health Professions Sci.				Humanities			
	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.
	Positive Count Component				Positive Count Component				Positive Count Component				Positive Count Component			
Gender is Male	−0.032		0.969	0.017	−0.047		0.954	0.025	0.132	***	1.141	0.015	0.082		1.085	0.054
Institution is Public	−0.180	***	0.835	0.015	0.040		1.041	0.033	−0.087	***	0.916	0.016	0.003		1.003	0.058
Institution is AAU Member	0.237	***	1.267	0.014	0.208	***	1.231	0.025	0.257	***	1.293	0.015	0.098		1.103	0.056
Years Since Degree	−0.004	***	0.996	0.001	0.005	**	1.005	0.002	−0.003	***	0.997	0.001	−0.001		0.999	0.003
Professorial Rank	−0.015		0.985	0.012	−0.125	***	0.883	0.022	0.040	**	1.041	0.013	0.002		1.002	0.049
Has Won Research Grants	0.333	***	1.395	0.014	0.496	***	1.642	0.030	0.343	***	1.409	0.016	0.390	***	1.476	0.091
Total Article Count	0.034	***	1.034	0.000	0.048	***	1.049	0.001	0.035	***	1.035	0.000	0.238	***	1.268	0.011
Log(θ)	0.672	***	1.959	0.019	0.349	***	1.417	0.037	1.071	***	2.919	0.023	−2.796	***	0.061	0.485
Constant	1.010	***	2.746	0.028	0.547	***	1.728	0.050	1.068	***	2.910	0.026	−3.974	***	0.019	0.507
	Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component			
Gender is Male	−0.097		0.908	0.055	−0.100	*	0.904	0.050	0.390	***	1.476	0.064	−0.047		0.954	0.033
Institution is Public	−0.183	***	0.833	0.050	−0.026		0.974	0.064	0.059		1.061	0.072	0.008		1.008	0.036
Institution is AAU Member	0.232	***	1.261	0.043	0.125	*	1.133	0.052	0.384	***	1.468	0.068	0.300	***	1.350	0.033
Years Since Degree	−0.014	***	0.987	0.002	−0.002		0.998	0.003	0.001		1.001	0.003	−0.012	***	0.988	0.002
Professorial Rank	−0.003		0.997	0.038	−0.022		0.978	0.044	0.029		1.029	0.048	0.041		1.042	0.031
Has Won Research Grants	0.545	***	1.725	0.050	0.561	***	1.752	0.110	0.674	***	1.962	0.110	0.389	***	1.475	0.072
Total Article Count	0.263	***	1.300	0.005	0.359	***	1.432	0.007	0.438	***	1.550	0.010	0.428	***	1.534	0.006
Constant	−0.540	***	0.583	0.085	−1.847	***	0.158	0.093	−1.827	***	0.161	0.097	−2.596	***	0.075	0.064
Observations	22,658				13,609				13,098				34,159			
Log-likelihood	−54,880				−21,960				−31,990				−20,370			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 8. Regression results from the Total OA Articles hurdle models for Natural Resources and Conservation, Physical and Mathematical Sciences, and Social and Behavioral Sciences, showing the count data component (Total OA articles > 0) and the zero count component (Total OA articles = 0)

	Nat. Res. & Conservation				Phys. & Math. Sci.				Soc. & Behav. Sci.			
	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.	β	Sig.	e^β	Std. Err.
	Positive Count Component				Positive Count Component				Positive Count Component			
Gender is Male	0.011		1.011	0.026	0.042	**	1.043	0.015	0.016		1.017	0.015
Institution is Public	−0.071	*	0.932	0.033	−0.099	***	0.906	0.014	−0.060	***	0.942	0.016
Institution is AAU Member	0.106	***	1.112	0.024	0.287	***	1.332	0.013	0.219	***	1.245	0.015
Years Since Degree	−0.004	**	0.996	0.001	−0.005	***	0.995	0.001	0.003	**	1.003	0.001
Professorial Rank	0.023		1.024	0.021	0.098	***	1.103	0.011	−0.079	***	0.924	0.013
Has Won Research Grants	0.235	***	1.265	0.024	0.350	***	1.419	0.013	0.341	***	1.407	0.017
Total Article Count	0.038	***	1.039	0.001	0.021	***	1.021	0.000	0.064	***	1.066	0.001
Log(θ)	1.369	***	3.932	0.042	0.209	***	1.232	0.015	0.581	***	1.788	0.024
Constant	1.141	***	3.131	0.051	1.222	***	3.393	0.026	0.391	***	1.478	0.029
	Zero Hurdle Component				Zero Hurdle Component				Zero Hurdle Component			
Gender is Male	0.040		1.041	0.139	0.070		1.072	0.052	0.111	***	1.118	0.033
Institution is Public	0.140		1.150	0.224	−0.078		0.925	0.050	−0.139	***	0.870	0.036
Institution is AAU Member	−0.060		0.942	0.139	0.388	***	1.474	0.046	0.428	***	1.534	0.033
Years Since Degree	−0.009		0.991	0.007	−0.016	***	0.984	0.002	−0.009	***	0.991	0.002
Professorial Rank	−0.128		0.880	0.110	0.090	*	1.094	0.039	0.000		1.000	0.030
Has Won Research Grants	0.408	*	1.504	0.181	0.427	***	1.533	0.053	0.476	***	1.610	0.055
Total Article Count	0.442	***	1.555	0.021	0.519	***	1.680	0.009	0.342	***	1.408	0.005
Constant	−0.870	**	0.419	0.296	−1.105	***	0.331	0.087	−1.469	***	0.230	0.060
Observations	3,913				30,415				28,228			
Log-likelihood	−9,894				−87,620				−51,720			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

scholars who have published at least one OA article) and the zero count component (predicting the binary variable "scholar has/has not published at least one OA article." The exponentiated coefficients for the truncated positive counts component are displayed in Figure 4.

4.3.1. Zero count component of hurdle model: Total OA articles

In all fields, the total number of articles authored was a positive and significant predictor of having authored at least one OA article, suggesting that increased overall publication activity increases the likelihood of at least some articles ultimately appearing in OA venues. For each additional article authored, the likelihood of having authored at least one OA article increases by a factor between 1.28 (Business) and 2.36 (Biological and Biomedical Sciences) (Tables 3–5). Securing federal research grants over the 5-year study period was a positive and significant predictor of authoring at least one article OA article in nine of 11 fields (all except Agricultural Sciences and Education). Many fields show a marked increase in the likelihood of authoring OA articles if federal research grants were won, including eight fields in which the likelihood of publishing at least one OA article increases by a factor more than 1.5 if a scholar has won a federal grant (Tables 3–5). Likewise, AAU membership was a consistently strong predictor of having published at least one OA article; researchers at AAU member institutions are more likely to have published at least one OA article than their non-AAU counterparts in nine fields (Biological and Biomedical Sciences; Business; Engineering; Family, Consumer and Human Sciences; Health Professions Sciences; Humanities; Physical and Mathematical Sciences; and Social and Behavioral Sciences).

Men are significantly more likely than women to have authored at least one OA article in three fields (Education, Health Professions Sciences, and Social and Behavioral Sciences), and women are significantly more likely than men to have authored at least one OA article in one field (Family, Consumer, and Human Sciences). Scholars at public institutions are significantly less likely to have authored at least one OA articles than their peers at private institutions in three fields (Business,

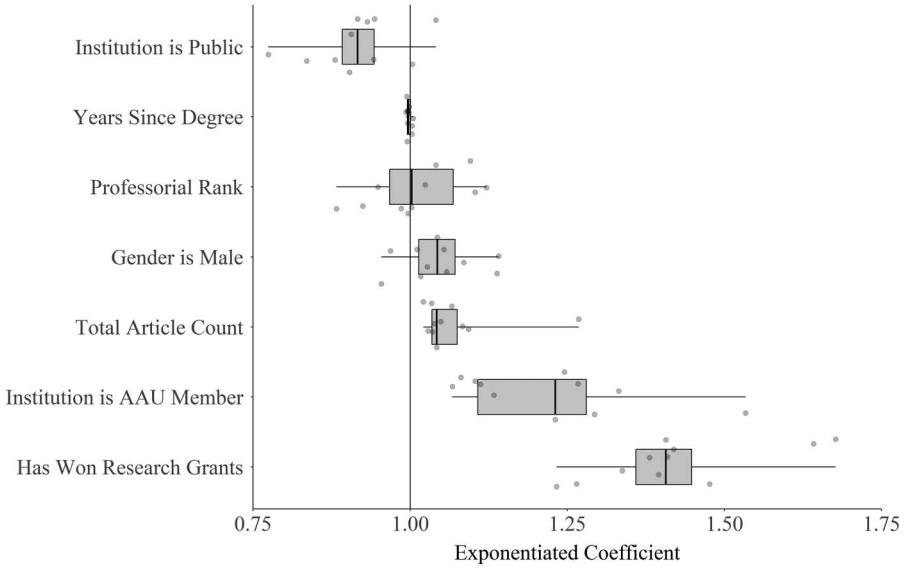

Figure 4. Range of exponentiated coefficients from the positive count component of the Total OA articles hurdle model for each independent variable. Values greater than the vertical reference line at 1.0 represent greater likelihood of publishing OA articles; values less than 1.0 represent lower likelihood of publishing OA articles. Individual dots represent the actual value for each of the 11 fields that constitute the box and whisker summary. The line in the center of each box is the median, the ends of boxes are the 25th and 75th percentiles, and the whiskers represent the full data range.

Engineering, and Social and Behavioral Sciences). Years since degree has a significant negative relationship with having authored at least one OA article in five fields (Business, Engineering, Humanities, Physical and Mathematical Sciences, and Social and Behavioral Sciences). Increased professorial rank (from assistant professor, to associate professor, to full professor) has a positive significant relationship with having authored at least one OA article in two fields (Business and Physical and Mathematical Sciences).

4.3.2. *Truncated positive count component of hurdle model: Total OA articles*

In every field, federal research grant support and the total count of articles authored were positive significant predictors of the number of OA articles published (Tables 6–8 and Figure 4). Having won at least one federal research grant results in an expected increase in the number of OA articles authored by a factor of between 1.23 (Education) to 1.68 (Business). Each additional article authored increases the likelihood of authoring an additional OA article by a factor of between 1.02 (Physical and Mathematical Sciences) and 1.27 (Humanities). A researcher's affiliation with an AAU institution also positively predicts the number of OA articles authored in all 11 fields (this is significant in nine fields) by a factor ranging between 1.08 (Agricultural Sciences) and 1.53 (Business).

Our model shows that scholars affiliated with public universities are expected to author fewer OA articles than their peers at private universities in seven of the 11 fields (Tables 6–8). Business and Engineering show the lowest exponentiated coefficient, with public institution scholars authoring OA articles at a factor of 0.77 and 0.84 that of their private institution colleagues, respectively. In two disciplines (Family, Consumer, and Human Sciences and Humanities) scholars at public institutions are expected to author slightly more OA articles than their peers at private institutions, although this effect is not significant in either case.

In eight of the 11 fields, men author more OA articles than women, and this result is significant in four fields (Biological and Biomedical Sciences, Education, Health Professions Sciences, and Physical and Mathematical Sciences). The strongest effect is in Health Professions Sciences, where men are expected to author more OA articles than women by a factor of 1.14 (Table 7). In two fields (Engineering; Family, Consumer, and Human Sciences), women are expected to publish more OA articles then men, although this effect is not significant in either field.

Years since terminal degree is a positive significant predictor of OA articles authored in two fields (Family, Consumer, and Human Sciences; Social and Behavioral Sciences), and a negative significant predictor in six fields (Tables 6–8). In all 11 fields, the exponentiated coefficient is between 0.99 and 1.01, suggesting an increase or decrease of 0.1 OA article authored for each additional year since terminal degree (Figure 4). Prediction of OA articles authored by professorial rank shows a relatively wide range of exponentiated coefficients, from 0.88 (Family, Consumer, and Human Sciences) to 1.12 (Biological and Biomedical Sciences). In two fields, increased professorial rank has a significant negative relationship with OA articles authored (Family, Consumer, and Human Sciences; Social and Behavioral Sciences), and in four fields increased rank has a positive relationship with OA articles authored (Biological and Biomedical Sciences, Business, Health Professions Sciences, and Physical and Mathematical Sciences).

5. DISCUSSION AND CONCLUSIONS

We found that 46.3% of the articles authored by the research professoriate in the United States over a 5-year period are OA articles (including all types of OA). This result shows close agreement with Piwowar et al.'s (2018) estimate based on Unpaywall data (46.7%–47.3%). The two regression

models we performed for each broad field of study (one predicting APC OA articles authored and one predicting Total OA article authored) yielded similar results. Both models reveal that federal research grant support is an especially strong predictor of OA publishing activity, suggesting that mandates by funding agencies to produce freely accessible research results are prompting greater levels of OA publishing. Both models also demonstrate that greater overall publishing activity predicts that at least some of the articles authored will be available as OA articles.

The model predicting APC OA articles published and the model predicting Total OA articles published differ with respect to the effect of gender on articles authored. In the APC OA model, we found that male gender predicts greater APC OA article authorship counts in 10 of 11 fields, including two fields (Education and Health Professions Sciences) where men are expected to author more than 1.3 times as many APC OA articles as women. The Total OA model also shows that men are more likely to have authored OA articles overall in most fields, but the factor by which men author more OA articles than women is lower in the Total OA model than in the APC OA model (cf. exponentiated coefficients in Figures 3 and 4). The gender gap in overall article output is well documented (e.g., Ceci et al., 2014), and our results appear to mirror findings about the gender gap in general: In most fields, men publish more OA articles than women, an effect that grows stronger when APCs must be paid. Duch et al. (2012) found that access to research resources (including grant support) and the relative career risks associated with different academic fields drive the gender gap in publication rates and impact. In their study, career risk for different academic fields was a function of "factors such as the time *T* to reach career independence, the fraction *A* of Ph.D. graduates that go on to careers in academia, and the reciprocal of the salary premium of nonacademic careers" (Duch et al., 2012, p. 7). Our data cover a different set of fields than Duch et al., precluding direct comparison using their model of career risk. Nonetheless, it is notable that the fields in our study with the greatest likelihood of men authoring more APC OA articles than women are Education, Health Professions Sciences, and Biological and Biomedical Sciences, which represent a diversity of average times to complete the Ph.D. degree (Bourke, Holbrook, et al., 2004) and offer several career paths outside the academy.

Our results show that each additional year since terminal degree (a proxy for academic age) predicts that a scholar will author as much as ±0.01 additional OA article (including APC OA articles). In STEM fields (Biological and Biomedical Sciences; Physical and Mathematical Sciences; Engineering), greater years since Ph.D. predicts fewer APC OA articles published, while in Business, Social and Behavioral Sciences, and Family, Consumer, and Human Sciences greater academic age predicts more APC OA articles published. The results for STEM fields appear to contradict the position that OA publishing may be perceived as risky by younger scholars seeking promotion or tenure (Agrawal, 2014), at least in those fields, suggesting that the perceived risk of publishing in OA venues may be limited to non-STEM areas, or possibly there is an alternative cause for field-level differences in the relationship between academic age and APC OA articles published.

Professorial rank also has a mixed relationship with (APC) OA articles published: Greater rank has a negative relationship with (APC) OA articles authored in some fields and a positive relationship in others. A positive relationship between rank and APC OA articles published was found in STEM fields (Biological and Biomedical Sciences, Engineering, and Physical and Mathematical Sciences), while a negative relationship was found in non-STEM fields (Business, Family, Consumer, and Human Sciences; Humanities; and Social and Behavioral Sciences). In 10 of 11 fields, years since terminal degree and professorial rank have the opposite relationship with APC OA articles published: In most STEM fields greater rank and fewer years since terminal degree predict greater APC OA articles, while in many non-STEM fields lesser rank and greater years since terminal degree predict greater APC OA articles. Taken together, these results may indicate that scholars who are promoted at a younger academic age are more likely to publish APC OA articles

than their equally ranked but older colleagues in STEM fields, with the opposite set of conditions describing most non-STEM fields. Nonetheless, field-level differences in APC OA publishing should be considered in future studies on OA publishing behavior and understanding field-level differences in OA authorship may be a fruitful direction for future research.

Analysis of the relationship between institutional characteristics and (APC) OA articles authored shows that scholars at AAU member institutions are more likely to author APC OA articles in most fields, and are more likely to author OA articles overall in every field. Scholars at private institutions are more likely to author APC OA articles in all but one field, and private institution scholars are more likely to author OA articles overall in all but two fields. The AAU (http://www.aau.edu) is an invitation-only group of 65 highly research-active North American research universities, membership is considered highly prestigious, and scholars at member institutions demonstrate higher rates of research publication and securing grants (Ali, Bhattacharyya, & Olejniczak, 2010). Private institutions in the United States are often associated with greater prestige than their public counterparts, evidenced by the consistently higher rankings of American private institutions in national and global ranking schema (e.g., *US News & World Report*; *Times Higher Education*). Clauset, Arbesman, and Larremore (2015) found that increased institutional prestige is related to increased faculty production, and Way, Morgan, et al. (2019) report that the prestige of a researcher's place of work, rather than the prestige of the doctoral program where they trained, has a positive effect on their scholarly productivity. Our data do not allow direct testing of doctoral program versus current workplace prestige, but our findings add "(APC) OA article authorships" to a growing list of scholarly productivity indicators that appear to increase with the prestige of a researcher's current institution.

Figure 1 reveals substantial field-level disparities in the proportion of APC OA articles authored by scholars in different fields. Scholars in three STEM fields (Biological and Biomedical Sciences, Physical and Mathematical Sciences, and Natural Resources and Conservation) author more than 25% of their articles in an APC OA venue (Gold and Hybrid). Engineering represents a conspicuous outlier among STEM fields, with only 19% of authorships occurring in APC OA venues (cf. Piwowar et al., 2018; Figure 4). Engineering falls between the range of other the other STEM fields and non-STEM fields; for example, Humanities scholars author only 8.1% of their articles in APC OA venues, and Social and Behavioral Science scholars author 13.1% of their articles in APC OA venues. In addition to the overarching trends, wherein men at more resource-rich institutions author more APC OA articles, field-level differences in the percentage of APC OA authorships must also be taken into account in studies of APC OA adoption and when crafting policies to further democratize the research literature.

The demographic and institutional characteristics we examined show that scholars at more prestigious universities (AAU members and private institutions), with greater federal research funding support, and who are willing to absorb greater career risk (men, other than the most highly qualified women; see discussion by Duch et al., 2012) tend to author more (APC) OA articles. Rates of participation in OA authorship are also greater in STEM fields, proportional to the total number of articles published (Figure 2). Among the goals of the rapidly growing OA movement is to make research results freely available to the public with no restrictions; our results indicate that the work of one subset of scholars (professors at institutions with greater resource availability, mainly in STEM fields) is disproportionately represented among this freely available portion of the literature. It has been suggested that OA articles and articles with greater "discoverability" (e.g., those available on websites such as Academia.edu) tend to be more highly cited than paywalled articles (e.g., Archambault, Côté, et al., 2016; Niyazov, Vogel, et al., 2016). Considering our results, the citation advantage of OA publishing likely benefits men at more elite institutions with greater research grant support.

The relatively new field of the science of science has provided valuable insights into the role of institutional prestige in knowledge production (Clauset et al., 2015; Way et al., 2019). As Way et al. note, research on the science of science sometimes falsely "…assumes, implicitly if not explicitly, that meritocratic principles or mechanisms govern the production of knowledge, p. 10733." The OA publishing model succeeds in democratizing the products of knowledge producers, but the knowledge producers whose work is published as OA articles are not necessarily representative of the broader research community. The disproportionately larger numbers of OA articles from professors at elite institutions represent a challenge to the OA business model: to increase the representation of scholars at a diversity of institutions backed by varying levels of research support among the OA literature.

ACKNOWLEDGMENTS

The authors wish to thank R. Berdahl, P. Lange, W. Savage, T. Stapleton, J. Sullivan, G. Walker, R. Wheeler, and C. Whitacre for valuable conversations and comments on drafts of this manuscript.

AUTHOR CONTRIBUTIONS

Anthony J. Olejniczak: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Writing—original draft, Writing—review & editing. Molly J. Wilson: Conceptualization, Data curation, Investigation, Methodology, Writing—original draft, Writing—review & editing.

COMPETING INTERESTS

Both authors are employed by Academic Analytics, LLC. Academic Analytics, LLC management had no oversight or involvement in the project and were not involved in preparation or review of the manuscript.

FUNDING INFORMATION

Data and computational resources were provided by Academic Analytics, LLC.

DATA AVAILABILITY STATEMENT

The data tables used to perform this study and supplementary materials are available via OSF (https://osf.io/sb8fq/) and have been assigned a DOI (10.17605/OSF.IO/SB8FQ). Records in the data set are anonymized to protect the privacy of individual scholars; the full (nonanonymized) data were offered to the editors and referees.

REFERENCES

Agrawal, A. A. (2014). Four more reasons to be skeptical of open-access publishing. *Trends in Plant Science*, *19*(3), 133. **DOI:** https://doi.org/10.1016/j.tplants.2014.01.005, **PMID:** 24521978

Ali, M., Bhattacharyya, P., & Olejniczak, A. (2010). The effects of scholarly productivity and institutional characteristics on the distribution of federal research grants. *The Journal of Higher Education*, *81*(2), 164–178. **DOI:** https://doi.org/10.1353/jhe.0.0084

Archambault, E., Amyot, D., Deschamps, P., Nicol, A., Rebout, L., & Roberge, G. (2014). *Proportion of Open Access papers published in peer-reviewed journals at the European and world levels: 1996–2013*. Brussels: European Commission. https://science-metrix.com/sites/default/files/science-metrix/publications/d_1.8_sm_ec_dg-rtd_proportion_oa_1996-2013_v11p.pdf

Archambault, É., Côté, G., Struck, B., & Voorons, M. (2016). *Research impact of paywalled versus open access papers*. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=https://www.bing.com/&httpsredir=1&article=1028&context=scholcom

Beaubien, S., & Eckard, M. (2014). Addressing faculty publishing concerns with open access journal quality indicators. *Journal of Librarianship and Scholarly Communication*, *2*(2), 1133. **DOI:** https://doi.org/10.7710/2162-3309.1133

Björk, B. C. (2019). Open access journal publishing in the Nordic countries. *Learned Publishing*, *32*(3), 227–236. **DOI:** https://doi.org/10.1002/leap.1231

Björk, B. C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Gudnason, G. (2010). Open access to the scientific journal literature:

Situation 2009. *PLOS ONE*, *5*(6). **DOI:** https://doi.org/10.1371/journal.pone.0011273, **PMID:** 20585653, **PMCID:** PMC2890572

Bourke, S., Holbrook, A., Lovat, T., & Farley, P. (2004). Attrition, completion and completion times of PhD candidates. *AARE Annual Conference*, *28*(January), 1–14.

Cameron, E. Z., White, A. M., & Gray, M. E. (2016). Solving the productivity and impact puzzle: Do men outperform women, or are metrics biased? *BioScience*, *66*(3), 245–252. **DOI:** https://doi.org/10.1093/biosci/biv173

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest, Supplement*, *15*(3), 75–141. **DOI:** https://doi.org/10.1177/1529100614541236, **PMID:** 26172066

Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, *1*(1), 1–7. **DOI:** https://doi.org/10.1126/sciadv.1400005, **PMID:** 26601125, **PMCID:** PMC4644075

Crotty, D. (2015). *Is it true that most open access journals do not charge an APC? Sort of. It depends*. The Scholarly Kitchen. http://scholarlykitchen.sspnet.org/2015/08/26/do-most-oa-journals-not-charge-an-apc-sort-of-it-depends/

Duch, J., Zeng, X. H. T., Sales-Pardo, M., Radicchi, F., Otis, S., … Nunes Amaral, L. A. (2012). The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLOS ONE*, *7*(12). **DOI:** https://doi.org/10.1371/journal.pone.0051332, **PMID:** 23251502, **PMCID:** PMC3520933

Fox, C. W., Burns, C. S., Muncy, A. D., & Meyer, J. A. (2016). Gender differences in patterns of authorship do not affect peer review outcomes at an ecology journal. *Functional Ecology*, *30*(1), 126–139. **DOI:** https://doi.org/10.1111/1365-2435.12587

Hine, C. (2010, June 13). Nebraska has it all to attract Big Ten, most importantly AAU membership. *The Chicago Tribune*. https://www.chicagotribune.com/sports/ct-xpm-2010-06-13-ct-spt-0614-aau-big-ten-expansion-20100613-story.html

Jackman, S. (2017). *{pscl}: Classes and Methods for {R} Developed in the Political Science Computational Laboratory*. https://github.com/atahk/pscl/

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., & West, J. D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius: Sociological Research for a Dynamic World*, *3*, 1–22. **DOI:** https://doi.org/10.1177/2378023117738903

Kleiber, C., & Zeileis, A. (2016). Visualizing count data regressions using rootograms. *The American Statistician*, *70*(3), 296–303. **DOI:** https://doi.org/10.1080/00031305.2016.1173590

Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B. C., & Hedlund, T. (2011). The development of open access journal publishing from 1993 to 2009. *PLOS ONE*, *6*(6). **DOI:** https://doi.org/10.1371/journal.pone.0020961, **PMID:** 21695139, **PMCID:** PMC3113847

Mabe, M. (2003). The growth and number of journals. *Serials: The Journal for the Serials Community*, *16*(2), 191–197. **DOI:** https://doi.org/10.1629/16191

Mishra, S., Fegley, B. D., Diesner, J., & Torvik, V. I. (2018). Self-citation is the hallmark of productive authors, of any gender. *PLOS ONE*, *13*(9), 1–21. **DOI:** https://doi.org/10.1371/journal.pone.0195773, **PMID:** 30256792, **PMCID:** PMC6157831

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, *33*(3), 341–365. **DOI:** https://doi.org/10.1016/0304-4076(86)90002-3

Niyazov, Y., Vogel, C., Price, R., Lund, B., Judd, D., … Shron, M. (2016). Open access meets discoverability: Citations to articles posted to Academia.edu. *PLOS ONE*. **DOI:** https://doi.or/10.1371/journal.pone.0148257, **PMID:** 26886730, **PMCID:** PMC4757559

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., … Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, *6*(2), 1–23. **DOI:** https://doi.org/10.7717/peerj.4375, **PMID:** 29456894, **PMCID:** PMC5815332

R Core Team. (2011). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. http://www.R-project.org/.

Siler, K., & Frenken, K. (2020). The pricing of open access journals: Diverse niches and sources of value in academic publishing. *Quantitative Science Studies*, *1*(1), 28–59. **DOI:** https://doi.org/10.1162/qss_a_00016

Solomon, D. J., & Björk, B.-C. (2012). A study of open access journals using article processing charges. *Journal of the American Society for Information Science and Technology*, *63*(8), 1485–1495. **DOI:** https://doi.org/10.1002/asi.22673

Stevenson, I. (2004). "Open access" for the rich only. *Times Higher Education*. https://www.timeshighereducation.com/comment/letters/open-access-for-the-rich-only/185910.article

Swan, A., & Brown, S. (2004). Authors and open access publishing. *Learned Publishing*, *17*(3), 219–224. **DOI:** https://doi.org/10.1087/095315104323159649

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. J. (2016). The academic, economic and societal impacts of Open Access: An evidence-based review. *F1000Research*, *5*, 1–55. **DOI:** https://doi.org/10.12688/f1000research.8460.1, **PMID:** 27158456, **PMCID:** PMC4837983

Way, S. F., Morgan, A. C., Larremore, D. B., & Clauset, A. (2019). Productivity, prominence, and the effects of academic environment. *Proceedings of the National Academy of Sciences*, *116*, 10729–10733. **DOI:** https://doi.org/10.1073/pnas.1817431116, **PMID:** 31036658, **PMCID:** PMC6561156

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, *27*(8), 1–25. **DOI:** https://doi.org/10.18637/jss.v027.i08