Heliyon 6 (2020) e05428

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

What drives citations of frontier application publications?

Justice G. Djokoto *, Kofi Aaron A-O. Agyei-Henaku, Akua A. Afrane-Arthur, Charlotte Badu-Prah, Ferguson K. Gidiglo, Francis Y. Srofenyoh

Agribusiness Management Department, Central University, Accra, Ghana

ARTICLE INFO

Keywords:
Social science
Citations
Efficiency
Frontier applications
Inverse hyperbolic sine
Ghana
Productivity

ABSTRACT

A large body of literature exists on analysis of citation and reviews of application of efficiency frontier. However, the reviews that assessed the determinants of citation counts did not focus on frontier applications. We contribute to the literature by identifying the drivers of citations of frontier application publications on Ghana. We employed two-part mixture modelling with inverse hyperbolic sine (IHS) transformation of the second part, which was found to be more appropriate than single equation IHS transformation modelling, for our data. Use of stochastic frontier analysis or data envelopment analysis did not drive citations counts. However, quality of journals in which frontier application studies were published and accessibility of the journals to readers, drive citation counts. Authors, institutions and funders of studies on frontier applications may consider these over collaborations, in seeking growth in citation counts.

1. Introduction

The inefficiency of a data management unit (DMU) is determined by measuring its distance to a frontier, indicating the potential for an efficiency increase (Lampe and Hilgers, 2015). The frontier shows the maximum outputs with different input combinations or the minimal combination of necessary inputs for different outputs. DMUs below or above the frontier are interpreted as inefficient and those DMUs on the frontier as efficient (Constantin et al., 2009; Lampe and Hilgers, 2015).

The methodologies of measurement of these efficiencies have led to two broad strands of approaches; data envelopment analysis and stochastic fronter analysis (Aigner et al., 1977; Banker et al., 1984; Charnes et al., 1978; Fare et al., 1983; Meeusen and van den Broeck, 1977). These frontier applications have culminated in a plethora of studies necessitating several reviews (Daraio et al., 2020; Lampe and Hilgers, 2015). The reviews covered more than half of the industries in the United Nations' International Standard Industrial Classification (ISIC) (Daraio et al., 2020; United Nations, 2008) and almost all the alphabetical categories of the Journal of Economic Literature (JEL) classification scheme (Cherrier, 2017; Daraio et al., 2020; Kosnik, 2018).

References to these frontier applications have resulted in many citation counts. Citations are the acknowledgements that one document receives from another (Bornmann et al., 2020; Garfield et al., 1979; Kaur and Rattan, 2018; Price, 1970). Citations of efficiency studies have varied from as high as 34,864 (Charnes et al., 1978) through 12,361 (Aigner et al., 1977) to as low as 0 (Yenibehit et al., 2019). As citations can be the explanandum (that to be explained) (Leydesdorff, 1998), we pose the following question: what factors drive or constrain the citations of references to frontier studies? Specifically, do collaborations, content, scope, accessibility and quality of the studies explain the observed variations in citations of frontier studies?

Lampe and Hilger (2015), Ahmad et al. (2020) and Liu et al. (2013) conducted citation-based reviews on efficiency studies globally. Using inverse hyperbolic sine (IHS) transformation to citations, Azoulay and Lynn (2020) studied the determinants of self-citation in the life sciences. Staudt (2018) studied the determinants of publication access in the National Institute of Health database. Hengel and Moon (2020) examined the role of authorship of females on citation impact in 'top five' economics journals. Whilst the citation-based reviews did not model the determinants of the citation counts, those studies that modelled the determinants of citation, did not attract the attention of studies on frontier applications. Indeed, we are unaware of any study that examined the drivers of citation of frontier studies. We fill these voids by identifying the drivers of citation of frontier publications with data on Ghana. Beyond the use of negative binomial, zero inflated Poisson regressions and single equation inverse hyperbolic (HIS) transformation of citation counts, we employed two-part mixture modelling with IHS transformation of the second part. Although our choice is rare in scientometrics, we found it to be more appropriate than single equation IHS

* Corresponding author.
E-mail address: dgameli2002@gmail.com (J.G. Djokoto).

https://doi.org/10.1016/j.heliyon.2020.e05428
Received 24 August 2020; Received in revised form 19 October 2020; Accepted 30 October 2020

transformation modelling, for our data, as shown in section three and four.

Identifying the drivers of citations in efficiency will inform authors, editors, funders, and policy makers as to the nature and extent of collaborations to forge and facilitate. These stakeholders may also desire to know the nature and extent of accessibility and quality of journals that would enhance visibility of published works on frontier applications. We, therefore, hypothesise as follows: i. collaborations do not influence citation counts of studies on frontier applications. ii. Content of frontier application studies does not affect their citation counts. iii. The scope of the frontier application study does not influence the citation counts. iv. The accessibility of the frontier application studies does not determine their citation counts. v. The quality of the destination of the publications does not explain counts of citation of studies on frontier applications.

The next section reviews the theoretical underpinnings of citation as well as the empirical evidence of drivers of citation counts. The literature search, data extraction and modelling are outlined in section 3. The result of the analyses is presented and same is discussed, encompassing implications of policy in section 4. Section 5 concludes.

2. Literature review

2.1. Theoretical review

Scientometricians have lamented the absence of an overarching theory of citing (Cozzens, 1981, 1985, 1989; Cronin, 1981, 1984; Luukkonen, 1991; Wouters, 1998). We therefore attempt to piece together the various ideas that have been proposed as theory for citations although there is still no accepted grand theory of citation (Small, 1999; Van Raan, 1998; Yue and Wilson, 2004).

Described variously as theory of citing (Mulkay, 1974), theory of citation behaviour (Gilbert Nigel, 1977) and citation theory (Cozzens, 1981), draws much from the sociology of science. The goal of these theories is largely to describe the citation behaviour of authors - why the author makes citations, and how these do or do not reflect the actual research of the citer and use of the literature (Wouters, 1998). We bifurcate the literature on the theory of citation into those of the early saints (60s to the end of the 70s) and of the latter saints (80s to date). The former are rewards, property, persuasive, Prince's 'success breeds success' and Small's persuasive theories. The latter include Latour's, Rousseau's, Leydesdorff's and Van Raan's theories.

Also known as the Mertonian theory, the rewards theory in the late 1960s and 70s, provided a sociological dimension of analyses of citation (Luukkonen, 1997). The theory states that the originality of the previous work of scientists is an institutional form of awarding rewards for their efforts and citations are a means for providing such recognition and reward (Merton, 1979). The first suggestion of a Mertonian interpretation of the meaning of citations was presented as early as 1965 by Kaplan (1965), quite soon after, the first citation index covering all major fields of science was published in 1964 (Luukkonen, 1997). Luukkonen (1997) further explained that it was not until in 1979 that Robert Merton acknowledged his authority to this understanding in print, in the foreword of the book by Garfield (1979).

Kaplan (1965) views citation practices largely as a social instrument for living with problems of property rights and priority claims. Ravetz (1971), who does not depart from this view markedly, noted that references constitute a method of protecting individual property right. Thus, citing the references constitutes splitting the property in the literary and providing a return to the owner of the property used, indicating the work was fruitful (Gilbert Nigel, 1977).

Gilbert Nigel (1977) put forward the tools of persuasion view. That a scientist who has obtained results believed to be true and important must persuade the scientific community to share the opinions of the value of the work (Liu, 1993). As a condition of persuasion, authors typically show how the results of their work represent an advance on previous research (Gilbert Nigel, 1977). This is accomplished by relating the findings to the current literature of the field of the study by providing evidence and argument to convince the audience that the work has not been vitiated by error (Gilbert Nigel, 1977). Further, that appropriate and adequate techniques and theories have been employed, and that alternative, contradictory hypotheses have been examined and rejected. The step-by-step argument and deduction contained in the paper is useful in attaining persuasion (Gilbert Nigel, 1977). This point is akin to the claim that articles become highly cited because they are judged to be of high quality (Cozzens, 1981). This quality is judged by a community of scientists, not the individual, which comes to a consensus on importance (Gilbert Nigel, 1977).

Within the general theory of cumulative advantage is "success breeds success" (Price, 1976, p. 305). The course of future citation successes is determined statistically by the history of the paper cited. To the econometrician, this can be modelled as distributed lags. That is, previous citations explain current citations. Thus, citation can be construed as generated by a pull mechanism from previous citations rather than from a push mechanism of the papers that do the citing (Cozzens, 1981; Price, 1976).

Citations are symbols of the citing authors; thus, referencing is a process of transforming published documents into symbols (Small, 1978). Small (1978) goes on to distinguish between own authors symbols and standard symbols. The former refers to author's own private symbols for certain ideas used, whilst the latter, having the same meaning for a community or group of scientists. In the case of the former, this could be references used once or twice. Since the latter has same meaning to the community, this is used often and by extension by many in the research community.

These theories outlined so far are not independent in their ideas, rather interrelated, sometimes semantically, leading Cozzens (1981) to classify them as normative interpretation, interpretive account and the symbolic perspective.

We now turn our attention to the views of the late saints. Latour (1987) notes that, references in articles are among the resources that are under the authors' command in their effort at trying to concretise the point being made and to lend support to the knowledge claims (Luukkonen, 1997). Other resources include the editors of the journals which publish the articles, the referees of the journals, and the research funds which finance the pieces of research (Luukkonen, 1997). This view draws on the perfunctory and rhetorical functions of citations within the scientific community (Leydesdorff, 1998; MacRoberts and MacRoberts, 1987).

The theory of Rousseau (1992, 1998) states that the reasons that authors cite are because of ethics, a matter of shorthand, persuasion, number of authors and recently, interestingness (Liu and Rousseau, 2013). Whilst borrowing from Gilbert Nigel (1977) regarding persuasion, Rousseau explained that persuasion arises when there is departure from issues of right and wrong. Moreover, the larger the authors freedom to choose from references, the more the need to persuade readers by use of citations. Adding that, the selection of the reference is not based on quality among other reasons. Rousseau (1992, 1998) on one hand and Cozzens (1981) and Gilbert Nigel (1977) on the other hand, seemed to differ on the dimension of quality driving citations.

The nature of citations is either explanandum that is, something that needs to be explained or explanans; the explanation (Hassan and Loebbecke, 2017; Leydesdorff, 1998). Citation analysis is regularly used as a tool or an indicator within an explanans (Leydesdorff, 1998). This view provides a multidimensional analysis or use of citations. Viewed within econometrics, citations could be caused or be the causal variable. Stated differently, citations could be the explained variable or the explanatory variable. Empirical examples are the use of citations of publications in journals as indicators of the journal's quality (explanans). However, quality of publication leads to citations as the explanandum (Lee et al., 2010; Bornmann et al., 2012).

Van Raan (1998) criticised the constructivist style in sociology of science concerning practices of citation. Using statistical arguments, Van

Raan (1998) showed that the strong limitations of any citation theory is the perspective of citer. Consequently, Van Raan (1998) emphasised that, citations should be conceived of as inherent properties of a publication. Keywords also have such inherent properties. There are empirically established relations between the citation domain and the word domain, thus, developing a model concerning citations only, may be unnecessary. This view, bibliometric chemistry, involves conceiving of citations as binding properties hence, these properties can be disentangled. This does not depart from the Leydesdorff's, view of citation as the explanandum. These thoughts notwithstanding, Scientometricians still await a grand theory of citation.

2.2. Empirical review

We present the flip side of the theoretical review, the empirical review, in order to provide a basis for the discussion section. The empirical literature relates to the variable categories derived from the theoretical review as well as from the existing empirical literature.

We focus on collaborations; author, institution and country levels first. Collaborative research projects may be, on average, more creative and important than individual papers (Nomaler et al., 2013). This is due to recombining resources from centres or persons within different country and traditions that is expected to generate more specific results (Nomaler et al., 2013). Breakthrough innovations do arise from recombining ideas that have earlier remained unlinked (Fleming, 2001; Nomaler et al., 2013).

Bornmann (2017), Ronda-Pupo and Katz (2017), Struck et al. (2018) and Thelwall and Maflahi (2019) found statistically significant positive effects of author collaborations on citation counts in Web of Science (WOS) and Scopus databases across disciplines over the period 1976 and 2012. The modelling included Ordinary Least Squares (OLS), negative binomial and zero-inflated Poisson regressions. Thelwall and Maflahi (2019) noted however, that there are significant cross-country and field dichotomy in the level of cooperation among researchers as well as how collaboration influences citation counts. On the contrary, Maz-Machado and Jimenez-Fanjul (2018) and Wongkhae et al. (2017) found negative association between number of authors and journal citation counts in journals with Thomson Reuters impact factor (TRIF).

Card and DellaVigna (2017) studied the editorial decision-making using anonymised submissions to four major economics journals. In matching papers to the publication records of authors and to later Google Scholar (GS) citations, they found among other results that the number of authors increased with citations counts. Card and DellaVigna (2017) transformed the citations using IHS to account for skewed citations arising from publications with no citations and rise in publication citations with the passage of time.

Bosquet and Combes (2013), studying citation of individual scholars in economics using Heckman 2-step procedure reported that number of authors constitute an important explanatory variable. This finding was explained thus. Several authors presented joint publications in conferences, seminars and workshops. More diffusion is generated as academics discuss their new findings with their peers. This is measured through citations. As the collaborating authors are also identified with institutions, Bornmann (2017) and Struck et al. (2018) have shown that institutional collaborations and citation counts are positively associated, in both WOS and Scopus databases. This was found in biomedical sciences and management fields with data between 16,000 and 128,000 across the globe.

For collaborations at country level, the works of Narin and collaborators appeared to be the first that addressed country collaboration and citation impact (Narin et al. 1991; Narin and Whitlow, 1990). Specifically, that the number of citations for international papers are higher than for one-country ones, and even twice as high as for one -country one-institution papers. Another facet has been the comparison between one-country collaborations and many-country collaborations (Frenken et al., 2005, 2010; Persson et al., 2004). Bornmann (2017), Leydesdorff et al. (2019) and Thelwall and Maflahi (2019) reported significant positive relationship between citations and many-country authorships. These results are similar across models such as OLS, negative binomial and zero-inflated Poisson, with citations measured as counts in WOS or normalised citations. Whilst the results of Struck et al. (2018) covered all fields excluding Arts and Humanities, Bornmann (2017) focused on biomedical sciences. From the existing literature, collaborations generally positively drive citation counts.

Regarding scope of efficiency and productivity studies, Daraio et al. (2020) recently showed that review studies on frontier applications cover 13 out of the 22 areas of economic activity as described by the ISIC scheme. They also showed that, the review studies represented categories A to I of the JEL classification scheme. Thus, although none of the studies on determinants of drivers of citation cited, used the JEL, Bosquet and Combes (2013) found a weak relationship between specialisation in fields of economics and individual citation score in GS. Nevertheless, we expect that the wider the scope of a study, the wider the spectrum of authors that will be interested in the publication, and the more knowledge diffusion that would occur, hence the more the citations will be.

For content of frontier applications; SFA or DEA, we did not find extant literature on their effects on citation scores. However, Lampe and Hilgers (2015) did point out that SFA is used largely in economics research whilst DEA is used largely in operations research. The foundation literature on DEA (Banker et al., 1984; Charnes et al., 1978; Fare et al., 1983) appear to have more citations than SFA (Aigner et al., 1977; Meeusen and van den Broeck, 1977). As the number of publications influence citations (Garfield, 1979; Kaplan, 1965; Bosquet and Combes, 2013) and it is the references that generate the citations, the effects of SFA/DEA on citations scores may well tilt in favour of DEA.

The accessibility category includes journal indexes, free-full-text (open access), inclusion in ResearchGate and Academica.edu sites. Batooli et al. (2017), in analysing major Clinical Medicine articles in Asian institutions, showed that publications deposited in ResearchGate increased exposure of scholars leading to increased citation count. In analysing data of 1,823 articles published by the authors from four different universities, Sababi et al. (2017) found a significant positive association between complete text availability and the citation count. Further, publisher's version of the archived manuscripts and post-print versions, received more citations than non-OA articles, and the difference in the citation counts of post-print manuscripts and publisher's version articles was not significant. These findings were within the domain of publications posted to ResearchGate.

There are differences in citations between articles posted to Academia.edu and other articles from similar journals (Niyazov et al., 2016). Using matching and regression analyses and a sample size of 31,216 and controlling for field, impact factor, and other variables, Niyazov et al. (2016) further found that the median impact factor journal uploaded to Academia.edu received 16% more citations after one year, 51% more citations after three years, and 69% after five years than a similar article not available online. Also, that articles posted to Academia.edu had 58% more citations than articles only posted to other online venues, such as personal and departmental home pages, after five years.

According to Staudt (2018), in 2008, the National Institutes of Health (NIH) mandated that the full text of NIH supported articles be made freely available on PubMed Central (PMC). Following this mandate, Staudt (2018) used 1 million NIH articles and many matched samples to examine how this "PMC mandate" affected access to the biomedical literature and nature of publishing in biomedicine. Staudt (2018) found modest effect on citations.

Struck et al. (2018) using data on the Sciences over 2008–2012 fitted to negative binomial, OLS and zero-inflated Poisson model, found that full-text availability of publications (open access) enhanced citation counts. This is expected as open access publications ensure the full-text is available without hindrance. Easy accessibility to reading the full-text will enhance the opportunity of citation of the publication.

Using standard mean comparisons and a negative-binomial regression model with several factors to account for the potential selection effects and quality bias, Wohlrabe and Birkmeier (2014) showed that articles available on the internet freely have indeed higher citation count in WOS, Research Papers in Economics (RePEc) and GS for the discipline of economics. Alkhawtani et al. (2020) also showed that open access publishing is related to an increased citation and download, in the field of radiology, based on analysis of 86 freely available and 414 paid-wall access articles from Europe, North America and Asia. Further, open access articles were found to be cited more often, downloaded and shared than subscription access articles.

Razumova and Kuznetsov (2019) analysed the impact of open access (OA) models on citation metrics. Using different research fields, they found higher values and growth rates of the citation metrics for OA reference groups over the paywall group. Abbasi et al. (2019) found similar results for library and information science (LIS) hybrid journals in Scopus for a period of ten years (2004–2013). Their conclusion was based on descriptive and analytical statistics. Further, Hubbard (2017) found that most studies that examined the OA citation advantage focused on some specific journals, areas of specialisation, countries or world output. Since local citation patterns may differ from these bigger patterns, the study examined the citation advantage of gold OA journal articles and found that OA physics journals showed a definite open access citation advantage, whereas other disciplines showed no difference or no open access citation advantage (Hubbard, 2017). In analysing OA journals from across all the 27 subject areas of the Scopus database in 2015, Dorta-González and Santana-Jiménez (2018) found that there were no generalisable citation advantage of OA at the journal level.

It can be observed from the OA-paywall journal comparison that generally, OA (full-text) availability of publications provide citation advantage across many disciplines over time and across different citation count measures. There are, however, rare occasions where AO publications did not differ from paywall publications. Certainly though, none of the paywall publications had citation advantage over OA, after controlling for all possible confounding factors.

Bornmann (2017), Elkins et al. (2010) and Lee et al. (2010) found positive association between quality of journal and citation count. These relationships were found from OLS regression and Spearman's correlation analysis. The findings of Elkins et al. (2010) are consistent across four quality indicators; Scopus, SCImago, Eigenfactor and TRIF. Using charts, Bozzo et al. (2017) and Cartwright and Savino (2009), reported statistically insignificant but positive relationship between quality and citation counts, for orthopaedics and ophthalmology research publications between 1999 and 2014. For all individual articles published in 2016 and 2017 in journals in plastic surgery (9,823), Asaad et al. (2019) studied their 2018 citations. Whilst the distribution of the citation was skewed to the right, with most articles having not more than 1 citation, a statistically significant but weak positive relationship was established between the number of citations and the journal impact factor. Nuti et al. (2015) using the Gini coefficient approach demonstrated statistically significant correlation between journal impact factor and citation count in cardiovascular journals. Whilst these empirical evidences concur with the position of Gilbert Nigel (1977) and Cozzens (1981), these are at variance with Rousseau (1992, 1998).

From the empirical review, different forms of collaborations, diverse modes of accessibility to publications and levels of quality standards positively influence citation counts. Whilst the studies reviewed covered many disciplines including a blanket one for all 27 disciplines in Scopus database, none of the studies focused on frontier applications that has been found useful in many facets of activity in the world. Our paper fills this void.

The methods of analyses have ranged from use of correlation coefficients (Spearman's and Pearson's) to single equations; OLS and negative binomial and finally mixture models; zero-inflated Poisson models with and/or without transformation of the citation counts and Heckman's 2-step procedure. Regarding the transformation of the dependent variable (citation), the transformations have involved natural logarithm as well as the IHS transformations. The transformations were necessitated by the skewness of the citations arising from many publications with no citations and the rise in publication citation counts with the passage of time. Whilst the IHS transformation seeks to approximate the normal distribution, it is unsurprising that such approximations would not always pass the normality test. We move beyond the single equation estimation to the two-part mixture model, with IHS transformation of the non-zero citation counts. This is another contribution to the scientometrics literature.

3. Method

3.1. Data

The total number of documents in the GS database, without any language restriction, is between 160 and 165 million (Orduña-Malea et al., 2015; Waltman, 2016). With 389 million records, GS is currently the most comprehensive academic search engine (Gusenbauer, 2019). Moreover, more than 90% of the primary literature for systematic literature review is accessible from GS (Yasin et al., 2020). Further, the use of GS permits the presentation of a wholistic picture of knowledge dissemination in economics and related fields (Bosquet and Combes, 2013). Thus, we searched GS as the main database. In addition, we sought to benefit from enhanced search capability by augmenting GS with other independent search engines as well as those embedded in publishers' websites (Gusenbauer and Haddaway, 2020). These include sources such as EBSCOhost, Scopus, AgEcon search, Wileyonlinelibrary, EmeraldInsight, among others. Following Djokoto et al. (2020), the following words and terms were used in the title line and as keywords: 'Efficiency' and 'Ghana', 'Data Envelopment Analysis' and 'Ghana', 'DEA' and 'Ghana', 'Stochastic Frontier Analysis' and 'Ghana', 'SFA' and 'Ghana', 'Frontier Analysis' and 'Ghana', 'inefficiency' and 'Ghana' and finally, 'Stochastic frontier' and 'Ghana'. The search ended at 18:00GMT on 31st December 2019. The reference list of literature retrieved were further searched. To be included in the data set, efficiency must have been measured by SFA or DEA. The exclusion criteria included 1. studies that did not report all the data used in the study and request for same was not received at the time of analysis of the data. 2. Duplicate studies. 3. Conference papers, thesis and working papers for which the journal published version was available. Applying these criteria and removal of duplicates yielded 196 publications (Supplementary file). The citations were the full citation counts as recorded in GS between 10:00GMT and 11:00GMT on 31st of March 2020. The inclusion of publications in journals indexed in WOS and Scopus databases were ascertained from the respective databases. Prior to checking by other authors, one author coded the data. We followed the meta-analysis of economics research-network (MAER-Net) protocol to collect the data and to improve transparency and the quality of quantitative reviews (Stanley et al., 2013). The details of the variables are presented in Table 1.

3.2. Modelling

Studies in citation analyses must necessarily follow rules like scientific research design. A clearly defined research objective(s) (Shengbo, 2015); the ascertainment of content-related variables, similarity in elements that constitute the sample (Shengbo, 2015); the dependent variable, citation frequency, may comprise more than one kind of citation (Peritz, 1992); the use of model-based methods which could accommodate larger numbers of variables and acknowledging the skewness of citation count distributions (Peritz, 1992). Having accounted for the first three in the introduction and the data sections, the last two inform the current section.

Beyond the citation counts is the factors that account for them. This is borne out of the overriding goal of sociology of science and technology and the main motive behind citation theories; to model the behaviour

Table 1. List of variables and their description.

Name	Label	Description
Dependent variables		
W_Citation	Number of citations per year	Number of citations per year counting from 1 year after publication.
CBin	CBin = 1, 0 otherwise	Binary rendition of W_Citation such that zero values are zero coded and values greater than zero are coded as 1.
CGzero	Citations	Exact values of citations
CGtzero	W_Citation>0	Exact values of W_Citation greater than 1.
Independent variables		
Collaboration on publications		
Authors	Author counts of publication	Number of authors on paper
Instit	Institutional affiliation count of authors of publication	
Country	Country counts of authors	Number of country designation of all others
Scope of publication within economics		
JEL	Number of three-digit Journal of Economic Literature classification of paper	Number of three-digit Journal of Economic Literature classification of publication.
Content (Dimension of efficiency)		
	SFA = 1 and 0 otherwise	Publication employed stochastic frontier analysis or Data envelopment analysis.
Accessibility of publications		
AllVersion	Total number of versions of paper available online as stated by Google Scholar.	Total number of versions of paper available online as stated by Google Scholar.
FT_2	FT_2 = 1, and 0 otherwise	Open access paper with full-text of publication freely available on Google Scholar
RG	RG = 1, and 0 otherwise	Publication available on ResearchGate platform
ADE	ADE = 1, and 0 otherwise	Publication available on Academia.edu platform
Quality of publications		
IF	IF = 1, and 0 otherwise	Publication in a journal with Thomson Reuter's impact factor
WOS	WOS = 1, and 0 otherwise	Publication in a journal indexed in web of science
Scopus	*Scopus* = 1, and 0 otherwise	Publication in a journal indexed in Scopus
Prefixes		
ihs	Inverse hyperbolic sine	Inverse hyperbolic sine transformation
ln	Natural logarithm	Natural logarithm transformation

behind citations. Leydesdorff (1998) refers to this as the explanandum. Hence, the variable at the heart of the citation analysis is the citation count. From the literature review, factors considered to influence the citation counts include collaborations, scope, content, accessibility, and quality. Consequently, we specify Eq. (1);

$$Citation = f(\text{collaborations, scope, content, accessibility and quality}) \quad (1)$$

The constituents of these categories of variables are also detailed in Table 1.

A close inspection of the data revealed that 29% of the observations of the dependent variable were true zeros (zero from the data generation process unlike zero from censoring). This proportion of true zeros is such that there are justifications for special treatment (Aïhounton and Henningsen, 2019; Belotti et al., 2015; Peritz, 1992; Shengbo, 2015). Handling such data using linear regression without taking cognisance of the many zeros could pose challenges regarding the difficulties in deriving error terms that are normal (Belotti et al., 2015; Carboni, 2012). We used two approaches in the literature to address the problem: inverse hyperbolic sine transformation of the dependent variable (Bellemare and Wichman, 2019; Burbidge et al., 1988; Lavin et al., 2017; MacKinnon and Magee, 1990; Pence, 2006) and two-part mixture models (Cragg, 1971; Ciminata et al., 2020; Duan et al., 1984; Lee et al., 2019; Mihaylova et al., 2011).

Inverse hyperbolic sine transformations have been suggested by Anscombe (1948; 1950) to normalise the *t*-distribution. The transformation has been used in citation studies (Azoulay and Lynn, 2020; Card and DellaVigna, 2017; Card et al., 2020; Hengel and Moon, 2020; Staudt, 2018). The IHS transformation (or archsinh) involves applying equation (2) to a variable y to get an outcome, say $\tilde{y}$:

$$\tilde{y} = \text{arcsinh}(y) = \ln\left(y + \sqrt{(y^2 + 1)}\right) \quad (2)$$

This transformation can relate to both dependent and independent variables in the following combinations; (i) linear–arcsinh, (ii) arcsinh–linear, (iii) arcsinh–linear with a dummy independent variable, and (iv) arcsinh–arcsinh specifications (Bellemare and Wichman, 2019). Following the transformation(s), the model can be estimated by OLS. We chose (iii) as our concern has to do with the dependent variable and some dummy variables as independent variables.

Recently, Bellemare and Wichman (2019) proposed elasticities for the explanatory variables when either the dependent or the independent variables are transformed. For our chosen option (iii), consider the estimable equation of the form

$$\tilde{y} = \propto + \beta x + \varepsilon \quad (3)$$

where x is a continuous variable. Following Bellemare and Wichman (2019), to recover y from the left-hand side of the equation of interest after estimating 3, one must apply the IHS transformation –the inverse of the IHS, or sinh – on both sides,

$$\tilde{y} = \sinh(\widehat{\propto} + \widehat{\beta}x + \widehat{\varepsilon}) \tag{4}$$

To recover $\vartheta_{yx} = \frac{\widehat{\partial y}}{\partial x} \cdot \frac{x}{y}$, that is the product of $\frac{\widehat{\partial y}}{\partial x} = \widehat{\beta}\cosh(\widehat{\propto} + \widehat{\beta}x + \widehat{\varepsilon})$ and $\frac{x}{y}$;

$$\vartheta_{yx} = \widehat{\beta}\cosh(\widehat{\propto} + \widehat{\beta}x + \widehat{\varepsilon}) \cdot \frac{x}{y} \tag{5}$$

The above can be restated as

$$\vartheta_{yx} = \widehat{\beta}\cosh(arcsinh(y)) \cdot \frac{x}{y} = \widehat{\beta}x \cdot \frac{\sqrt{y^2+1}}{y} \tag{6}$$

In the limit $\lim_{y\to\infty} \frac{\sqrt{y^2+1}}{y} = 1$, for large values of y, so that $\vartheta_{yx} \approx \widehat{\beta}x$.

For dummy independent variables, Eq. (6) will not suffice. We followed Bellemare and Wichman (2019) to obtain the elasticities for the dummy variables. Consider the estimable equation

$$\tilde{y} = \propto + \beta d + \varepsilon \tag{7}$$

where d is a binary variable taking on only values of 0 or 1. The quantity $\partial\tilde{y}/\partial d$

is undefined because of the discrete changes in the binary variables. According to Bellemare and Wichman (2019), $\overline{P}$ as the percentage change in y associated with a change from $d = 0$ to $d = 1$ can be found as

$$\frac{\widehat{\overline{P}}}{100} = \frac{\widehat{y}(d=1) - \widehat{y}(d=0)}{\widehat{y}(d=0)} = \frac{\sinh(\widehat{\propto} + \widehat{\beta} + \widehat{\varepsilon}) - \sinh(\widehat{\propto} + \widehat{\varepsilon})}{\sinh(\widehat{\propto} + \widehat{\varepsilon})} = \frac{\sinh(\widehat{\propto} + \widehat{\beta} + \widehat{\varepsilon})}{\sinh(\widehat{\propto} + \widehat{\varepsilon})} - 1 \tag{8}$$

Following Halvorsen and Palmquist (1980), Eq. (8) yields

$$\frac{\widehat{\overline{P}}}{100} \approx \exp(\widehat{\beta}) - 1 \tag{9}$$

The two-part model has been employed in fields such as health (Dean et al., 2020; Lartey et al., 2020) and agriculture (Dehghan et al., 2020; Ruhinduka et al., 2020). Gutiérrez et al. (2020), Ramalho and da Silva (2009) and Wulff (2019) used two-part models to model DEA scores in aquaculture and dairy respectively. The two-part model requires fitting a binary choice model for the probability of observing a positive-versus-zero outcome. Subsequently, based on a positive outcome, a regression model is estimated for the positive outcome (Belotti et al., 2015; Ciminata et al., 2020; Zhao et al., 2020). Following Lee et al. (2019), let

$$\mathrm{C}_i = \left\{ \begin{array}{l} 0 \; if \; Y_i = 0 \\ 1 \; if \; Y_i > 0 \end{array} \right\} \tag{10}$$

$$Probit\{\Pr(\mathbb{C}_i = 1 | X_i, U_i)\} = \theta X_i + U_i \tag{11}$$

$$g(Y_i) | Y_i > 0, X_i = \beta X_i + V_i + \varepsilon_i \tag{12}$$

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right) \tag{13}$$

In equation (10), Y represents the value of the explained variable observed for citation of the publication and whether the response was nonzero or not is indicated by $\mathbb{C}$. Eq. (11) shows $\mathbb{C}$ modelled with probit regression with a random intercept U_i that accounts for systematic differences between publications (Belotti et al., 2015). The influence of driver and constraint variables, $\boldsymbol{X_i}$, is estimated by θ (Belotti et al., 2015). Eq. (12) shows how the magnitude of Y_i, conditional on it being greater than zero, depends on the driver variables $\boldsymbol{X_j}$ (Belotti et al., 2015). $g(\cdot)$ is a transformation function that produces a normal distribution of the outcome Y_j (Belotti et al., 2015). We modelled this as an IHS transformation variable. The influence of the variables, $\boldsymbol{X_j}$, are obtained as β and a random intercept V_j, which captures the systematic differences between publications (Farewell et al., 2017). It is worth noting that i derives from the total sample whilst j is the number of observations for which Y > 0. Therefore, logically, i - j > 0. The variables, Z, Y and X are detailed in Table 1.

4. Result and discussions

4.1. Background of the data

For the binary variable *CBin*, 71% of the publications recorded citations greater than zero (Table 2). Thus, 29% (196 - 140 = 56) of the 196 publications recorded true zero citations. This informed the consideration of the IHS transformation. An examination of the citations over time showed that half (28) of the 56 zeros occurred for the years 2017–2019. Also, 20 were for the period 2013 to 2016. This suggest that the shorter the publication is in circulation, the less likely it will be cited (Bornmann and Daniel, 2006). Thus, we weighted the citation counts by the length of years in circulation, hence *W_Citation*. The mean of *W_Citation* is 2.60 lower than that of CGtzero (3.64) because the latter is purged of publications with zero valued weighted citations.

The number of authors per publication varied from 1 (66 publications) to 9 (Asante et al., 2013). That for affiliations peaked at 8 and country collaborations peaked at 11. The differences in the peaks arose from some authors having more than one institutional affiliation and consequently country affiliation. There were as many as 11 publications with two-digit JEL classifications. Publications that used SFA were 75% of the sample. This is because most of the publications were in agricultural production, which is inherently stochastic hence the use of SFA.

There were as many as 23 different versions of a publication in GS. As many as 88% of the 196 publications possessed full-text in GS. Whilst 68% of the publications can be found in ResearchGate, a lower proportion, 28% have been posted in Academia.edu. Thus, ResearchGate appeared to have been more attractive with frontier researchers than Academia.edu.

Regarding the quality indicators, 16% of the publications were in journals indexed in Scopus. A lower percentage of 11 is indexed in WOS and less than 10% of the publications have found a home in journals with TRIF.

4.2. Estimations

In model 1, the *W_Citation* was transformed using IHS, as is the practice with citation data (Table 3). However, the Shapiro_Wilk W test has p value of 0.05 suggesting non-normality of the errors, a violation of the normality assumption of OLS estimation procedure. This necessitated the use of the two-part mixture model. The F-test of the binary model, model 2, is significant at p value less than zero. However, WOS and TRIF were dropped due to perfect fit. Since the zeros in the *W_Citation* are no longer part of the *CGzero*, model 3, the other part of the two-part model, was estimated using OLS. All OLS assumptions were not violated except that of normality as the Shapiro_Wilk W test for normality with p value of less than 0%. We performed IHS transformation on the *CGzero* to become *CGtzero* and estimated model 4.

The sign for WOS is negative and the coefficient is significant (model 4). This is unexpected as the literature suggests that the sign should be positive. Moreover, the other counterpart quality covariates are positively signed, appropriately. We sought to find the reason for this unexpected outcome, with a suspicion that collinearity could be the culprit. This was influenced by the literature that collinearity could result in estimates of the "incorrect sign" among other problems (Belsley et al.,

Table 2. Descriptive statistics.

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
Dependent variables					
W_Citation	196	2.5988	3.4069	0	16
CBin	196	0.7143	0.4529	0	1
CGtzero	140	3.6381	3.5315	0.08	16
Independent variables					
Collaboration					
Authors	196	2.3367	1.2886	1	9
Instit	196	1.4745	0.8619	1	8
Country	196	1.2806	0.9102	1	11
Scope					
JEL	196	4.5102	2.4062	1	11
Content					
SFA1	196	0.7551	0.4311	0	1
Access to the content					
AllVersion	196	5.2704	3.9695	1	23
FT_2	196	0.8776	0.3286	0	1
RG	196	0.6786	0.4682	0	1
ADE	196	0.2755	0.4479	0	1
Quality					
IF	196	0.0867	0.2822	0	1
WOS	196	0.1122	0.3165	0	1
Scopus	196	0.1634	0.3706	0	1

Table 3. Estimations of elasticities.

VARIABLES	(1)	(2)	(3)	(4)	(5)	Equality
	ihs_W_Citation	Cbin	CGtzero	ihs_CGtzero	ihs_CGtzero	(2)–(5)
Collaboration						
Authors	0.2430 (0.1549)	0.0248 (0.0348)	0.1282 (0.2106)	0.2272 (0.1523)	0.2339 (0.1545)	-1.3203
Instit	0.1340 (0.1623)	0.0562 (0.0754)	0.3204 (0.4888)	0.0609 (0.2011)	0.0828 (0.2037)	-0.1225
Country	-0.1486 (0.1124)	-0.2679** (0.1099)	0.8566 (0.6784)	0.1470 (0.2455)	0.1028 (0.2482)	-1.3657*
Scope						
JEL	-0.0019 (0.1343)	0.0026 (0.0138)	0.0845 (0.1458)	-0.0301 (0.1452)	-0.0140 (0.1471)	0.1124
Content						
SFA1	0.9808*** (0.1518)	0.0488 (0.0779)	-1.2838 (0.8443)	-0.8869*** (0.1567)	-0.8895 (0.1594)	-4.7386***
Accessibility						
AllVersion	0.4022*** (0.1074)	0.0329*** (0.0112)	0.1350 (0.0865)	0.2421 (0.1056)	0.2227** (0.1068)	-1.7675**
FT_2	1.0964*** (0.2193)	0.0250 (0.1024)	0.1596 (0.7984)	1.1370*** (0.2507)	1.1032*** (0.2463)	-4.0422***
RG	1.6387*** (0.2427)	0.1366** (0.0685)	1.4566** (0.6153)	1.6411*** (0.2773)	1.6601*** (0.2484)	-5.9126***
ADE	1.1953*** (0.1836)	0.1090 (0.0832)	-0.1521 (0.5328)	1.0277*** (0.1582)	1.0641*** (0.1653)	-5.1611***
Quality						
IF	1.3585*** (0.4495)	-	3.6194** (1.4956)	2.1744** (0.9289)	1.1734*** (0.3807)	-
WOS	-	-	-3.4753** (1.4980)	-0.3623** (0.1688)	-	-
Scopus	1.5984*** (0.4111)	0.1183 (0.5208)	1.2848 (1.1685)	1.9012*** (0.5917)	1.3023*** (0.3409)	-1.9022**
Model properties						
Observations	196	174	140	140	140	-
R-squared	0.3672	-	0.2963	0.3226	0.2973	-
F test	9.70***	37.86***	9.25***	5.04***	4.92***	-
Bruesch-Pagan/Cook-Weisberg	2.34	-	-	0.26	0.09	-
Ovtest	1.39	-	0.28	1.62	1.16	-
Shapiro_Wilk W test	1.798**	-	5.814***	1.261	-0.735	-
Highest vif	2.27	-	6.20	6.20	3.26	-

1980; Greene, 1993; O' Brien, 2007; Wooldridge, 2009). First, the highest VIF for the second part of the two-part model was recorded by WOS as 6.20. Second, from the correlation matrix of the covariates, the coefficient (rho) of WOS/TRIF is 0.8610 whilst that for WOS/Scopus is 0.8268. Incidentally, for TRIF/Scopus, the rho is lower at 0.7119. The highest rho value for WOS/TRIF is because inclusion in Web of Science

database is a primary condition for a journal to obtain an TRIF. Thus, every TRIF journal is included in WOS. In our data, majority of the Scopus indexed journals are also in WOS. The computed levels of rho values are due to the similarity of the variables; all three variables measure quality of the journal. We do not seek to commit *partialling fallacy* (Gordon, 1968), rather we included these because they measure quality at different levels. Therefore, combining the VIF for WOS and the rho values of WOS with the others; TRIF and Scopus, we conclude that WOS is problematic hence, should be dropped from the covariates. Consequently, we estimated model 5. We consider the loss of the explanatory power of the model (Adjusted R squared) of 2.17% points to be negligible. Moreover, the *p* values of the coefficients that were below 0.05 remained below 0.05 with some moving to below 0.01. Thus, the second part of the model has one more covariate than the first. This is to be expected because there may be legitimate theoretical or statistical reasons for using different covariates in the two parts of the modelling (Belotti et al., 2015). It is worth noting that the model properties fulfil all the tested assumptions of the OLS.

The statistical significance of the F-tests (model 2 and 5) means the explanatory variables jointly explain the variability in the citation counts; whether measured as binary or as linear (Table 3). This confirms the underlying idea about theories of citation; explaining the behaviour of authors regarding citation. Leydesdorff (1998) specifically notes that the citation is the explanandum (that to be explained). The statistical significance of the F-tests thus, paves way for discussion of the coefficients of model 2 and 5.

A test for the equality of parameters for the first and second part of the two-part model shows that all parameters differ between the two models except for *Authors*, *Instit*, and *JEL*. The coefficients of *Authors*, *Instit* are positively signed. This agrees with existing studies (Bornmann, 2017; Card and DellaVigna, 2017; Ronda-Pupo and Katz, 2017; Struck et al., 2018; Thelwall and Maflahi, 2019). The sign, however, diverges with the findings of others (Maz-Machado and Jimenez-Fanjul, 2018; Wongkhae et al., 2017). On the other hand, the magnitude of the marginal effects and elasticities are not statistically significant. This means that although there is much to be gained through collaborations such as skills, ideas, new knowledge and reinforcement of existing knowledge and capabilities (Fleming, 2001; Nomaler et al., 2013), in the case of our data, these are inadequate to significantly induce increases in citation counts. This departs from the existing literature (Bornmann, 2017; Card and Della-Vigna, 2017; Ronda-Pupo and Katz, 2017; Struck et al., 2018; Thelwall and Maflahi, 2019). The findings suggest authors of frontier studies should seek other avenues of driving citations other than collaboration among authors, in the same institutions or other institutions.

Whilst the elasticity of *Country* for model 5 is positive and statistically insignificant, model 2 reports negative and statistically insignificant magnitude of the marginal effect. The differences in magnitudes of the accelerations have a *p* value of more than 0.10, higher than the cut-off of 0.05. Thus, not only are the accelerations statistically insignificant, the differences are also statistically insignificant. These findings diverge with those of existing studies (Frenken et al., 2010; Narin et al., 1991; Persson et al., 2004). As in the case of the earlier forms of collaboration, teamwork at the country level could generate benefits, but not enough to increase citation within the context of this study. Funders and authors may pursue collaborations for reasons other than to increase citations for publications relating to frontier applications.

The scope of frontier application within economics and content of the frontier application vis-à-vis SFA and DEA did not significantly influence citations counts. Wide scope of frontier application studies did not necessarily induce higher citations than those with narrow scope. This finding goes contrary to expectation. In the case of *SFA* whilst the accelerations of the variables are statistically insignificant the difference is statistically significant. The results imply that the choice between SFA and DEA in frontier studies would not increase citations.

DEA methods accommodate multiple outputs and inputs and they are also flexible with no requirement for imposing a functional form. DEA uses linear programming. Whilst it is non-parametric, the model is deterministic (noise is included in the efficiency score rather than accounted for directly). On the other hand, SFA is a stochastic model and requires specification of a functional form. It is a parametric model. Although SFA requires imposition of functional forms, there is a wide range of functional forms to choose from including the linear one. In the received literature, SFA accommodates single input (output) and multiple output (input). The estimation procedure is regression and mostly by maximum likelihood (Aigner et al., 1977; Banker et al., 1984; Charnes et al., 1978; Fare et al., 1983; Lampe and Hilger, 2015; Meeusen and van den Broeck, 1977). Frontier researchers may well continue to be motivated by the above properties of SFA and DEA in the choice of either SFA or DEA, as there are no significant gains to be made in citation counts for choosing between the two.

Two popular academic social media platforms are ResearchGate and Academia.edu. Full-texts of both open access and non-open access documents can be posted on them. However, the latter can only be shared privately between the poster and the requester. In this way, members of the community can share in both types of documents.

The marginal effect for Academia.edu (ADE) is statistically insignificant (model 2) whilst the elasticity of ADE is statistically significant in model 5. The statistical significance of the difference between two coefficients imply that, whilst posting documents in ADE does not significantly increase the probability of obtaining a citation, once cited, a document posted on ADE can increase citation counts by 1.1%. This finding is consistent with Niyazov et al. (2016). The inclusion of frontier study in RG would increase citation by 1.7%. The accelerations for both models for *RG* are statistically significant at *p* value less than 0.00, just as the difference between the estimates. Whilst the statistically significant magnitude and sign, are in accordance with existing literature (Batooli et al., 2017; Sababi et al., 2017), the significant difference in the parameters imply that the propensity to increase citation linearly is higher than the propensity to move citation from 0 to 1 and beyond. There appears to be some inertia to cite frontier papers in ADE.

The differences in the parameters of the variables can be explained by the higher patronage of ResearchGate (68%) than Academia.edu (28%) in our data. This explanation is buttressed by Bosquet and Combes (2013), that is the number of publications is positively associated with the citation counts. Also, recently, the latter has gone commercial whilst the former still has all its services for which payment is required in ADE, accessible at no cost to the user. Despite providing some services not-for-free, posting publications and downloading which are essential for citation are still offered at no cost to the users of ADE. Although the effects of the RG variable on the explained variable seemed stronger than that of the *ADE*, frontier researchers who are not currently in the two communities should register and post their publications. Others who are registered but do not have all their documents uploaded should do so. Patronising the two will yield benefits from both worlds as this could yield up to 2.8% rise in citations. Both platforms do rank institutions and countries by citations within them. Thus, by encouraging authors to build profiles and upload publications, institutions and funders can increase their exposure.

The statistical insignificance of *FT_2* (model 2) implies that although availability of full-text of publication in GS could increase the probability of being cited, the 2.5% likelihood is inadequate statistically. In the case of the *FT_2* (model 5), inclusion of full-text publication in GS will increase citation by 1.1% once the document obtains a citation. The positive and statistically significant result for frontier applications in Ghana, is consistent with that for diverse disciplines, countries and databases (Abbasi et al., 2019; Alkhawtani et al., 2020; Hubbard, 2017; Razumova and Kuznetsov, 2019; Staudt, 2018; Struck et al., 2018; Wohlrabe and Birkmeier, 2014). This finding underscores the actions of the open access movement.

Gilbert Nigel (1977) theorised that as a condition of persuasion, authors tend to show how their results represent progress on previous research. This is accomplished by relating the findings to the literature of

the field of the study by providing tangible outcomes and argument to convince the audience that the publication is not impaired by error. Gilbert Nigel (1977) further noted that the author will seek to show that "appropriate and adequate techniques and theories have been employed, and that alternative, contradictory hypotheses have been examined and rejected" (Pg., 116). Persuasion is attained by systematic argument and deduction detailed within the paper. The audience of the author can certainly not ascertain all these as noted by Gilbert Nigel (1977) using an abstract. Rather, a full-text would afford the reader the opportunity to effectively do so. Whilst the author will have opportunity to be cited with the complete information available in the full-text, the reader has ample opportunity to make that decision.

Whether gold or green, open access offers full-text availability to readers immediately after publication and opportunity for authors, institutions and funders to make the publications freely available to readers. The open access thus increases the exposure of the publication thereby increasing the opportunity to persuade readers to cite it (Gilbert Nigel, 1977). Following from this finding, authors of frontier studies who seek to gain citation of their works should look out for publication outlets/destinations that offer open access. As the difference between gold or green open access has to do with authors making payments prior to publication, authors should explore opportunities for full or partial waivers on article processing charges (APC). Some publishers also form partnerships with institutions with minimal or no fee payment such as Research4Life to make full-text publication freely available. Institutions that fund journals should continue to do so to ensure wider circulation and consequent citation.

As GS crawls the web to search for documents requested by the surfer, the codes can capture all available documents on the web. GS states one version with the full-text, freely and downloadable version on the right-side of the page with a hyperlink if there is one. Below the abstract, all versions of the publication found including that which is exhibited are stated as 'All…versions'. These include whether the publication is full-text, freely available or behind a paywall. The versions offer alternative opportunities to the surfer as well as increased exposure to the publication, which is beneficial to the author, institutional affiliation and funders. The *AllVersion* captures publications posted in ResearchGate and Academia.edu and others. These principally are derived from the indexing of the journals in which the publications find home. Increasing number of versions of publication in GS by 1% will increase the probability of citation by 3.2%. On the other hand, increasing all versions by 1% will increase citations greater than zero by 0.22%. This finding suggests databases and locations including ResearchGate and Academia.edu do provide exposure to publications and consequently citations. Authors seeking citation for their work could consider journals that have many indexing avenues. Additional exposures can be derived from archiving full-text at other locations accessible to GS.

Since the TRIF predicted success perfectly in the binary model, it was dropped. Thus, we have estimates for only the citations greater than zero. The elasticity shows that publication of a frontier application paper in a journal with TRIF would increase citations beyond zero by 1.17% above the reference, that is, those that are not published in journals with TRIF. This estimate is statistically significant at less than 0.00%, consistent with the existing literature (Asaad et al., 2019; Bornmann, 2017; Bozzo et al., 2017; Cartwright and Savino, 2009; Elkins et al., 2010; Lee et al., 2010).

The other quality variable, Scopus, also resulted in a positive and statistically significant elasticity of 1.3023. This implies publishing in a journal indexed in Scopus will increase citation of frontier application paper by 1.30 above those journals that are not indexed in Scopus. Whilst our results disagree with the theoretical position of Rousseau (1992, 1998), that selection of references does not depend on quality of the publication, it agrees with the persuasion theory of Gilbert Nigel (1977) and the position of Cozzens (1981) that articles become highly cited because they are judged to be of high quality. Our finding is also in line with the existing empirics (Asaad et al., 2019; Elkins et al., 2010; Nuti et al., 2015). The concordance of the quality variables shows that publishing in quality journals begets citations in GS. Indeed, journals of high quality may have been maintaining a more stringent gate-keeping role than others. In the most recent citation theory, Liu and Rousseau (2013) noted that interestingness could be another reason for citation. This may well be a characteristic of publications that find a home in journals of high quality. From the foregoing, authors seeking to maximise citations should consider publishing in journals of high quality, as a criterion. Aside of a personal decision for seeking citations, and where number of publications count towards promotion, institutions could consider publishing in journals with TRIF, a condition for tenure. Authors that currently do not publish in journals indexed in Scopus, WOS or with TRIF, should consider publishing in Scopus indexed journals as a first step towards publishing in journals traditionally regarded as being of a quality status. For those who currently publish in journals indexed in Scopus, moving up to journals with TRIF could provide additional increase in citation to their publications in GS.

It can be observed that the results of our study confirm most of the theoretical positions on citations. Also, the results are consistent with existing literature across countries, disciplines, time and other sources of citations. This consistency enhances the applicability of our findings beyond the frontier studies to other disciplines and from the national scope to regional and the global level.

5. Concluding remarks

Extant reviews did not assess the determinants of the citation counts, those studies that focused on determinants of citation, did not employ data on frontier applications. Indeed, we are unaware of any study that examined the drivers of citation of frontier application studies. We contribute to the literature by identifying the drivers of citations of frontier application publications. Specifically, for the country Ghana, this study provides another perspective of citation analysis to the existing bibliometric studies on Ghana. Beyond the use of single equation IHS transformation of citation counts, which appears to be a standard approach, we employed two-part mixture modelling with IHS transformation of the second part, which we found to be more appropriate than single equation IHS transformation modelling, for our data.

Author, institutional affiliations and country collaborations did not have discernible effect on citation counts for studies on frontier application. Coverage of study in terms of JEL classification as well as whether SFA or DEA, did not drive citation counts. All proxies for accessibility of publication (versions in GS, open access-full-text, posting in ResearchGate and Academia.edu) drive citations of frontier application studies. Finally, publishing in journals with TRIF and journals indexed in Scopus influenced citation counts of frontier publications.

Authors, funders and institutions seeking to increase citations to their publications on frontier applications should consider journals with the widest possible indexing, journals with open access for publications and journals of high quality. Notwithstanding the consistency of our findings with studies that cover other geographic areas and the globe, citation analysis of frontier applications' data at the regional and global levels will be worth exploring. The outcome would provide perhaps, a double assurance.

Declarations

Author contribution statement

J. G. Djokoto: Conceived and designed the experiment; Performed the experiments; Analysed and interpreted the data; Wrote the paper.

K. A. A-O. Agyei-Henaku, A. A. Afrane-Arthur, C. Badu-Prah, F. K. Gidiglo, F. Y. Srofenyoh: Performed the experiments; Contributed reagents, materials, analysis tools; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2020.e05428.

References

Abbasi, Z., Shekofteh, M., Shahbodaghi, A., Kazemi, E., 2019. Citation indicators' comparison of LIS open access and subscription publications based on Scopus. Global Knowledge, Memory and Communication 68, 288–299.
Ahmad, N., Naveed, A., Ahmad, S., Butt, I., 2020. Banking sector performance, profitability, and efficiency: a citation-based systematic literature review. J. Econ. Surv. 34, 185–218.
Aigner, D., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic production function models. J. Econom. 6, 21–37.
Aïhounton, G.B., Henningsen, A., 2019. Units of Measurement and the Inverse Hyperbolic Sine Transformation (No. 2019/10). IFRO Working Paper.
Alkhawtani, R.H., Kwee, T.C., Kwee, R.M., 2020. Citation advantage for open access articles in European Radiology. Eur. Radiol. 30, 482–486.
Anscombe, F.J., 1948. The transformation of Poisson, binomial and negative-binomial data. Biometrika 35, 246–254.
Anscombe, F.J., 1950. Table of the hyperbolic transformation. J. Roy. Stat. Soc. 113, 228–229.
Asaad, M., Kallarackal, A.P., Meaike, J., Rajesh, A., de Azevedo, R.U., Tran, N.V., 2019. Citation skew in plastic surgery journals: does the journal impact factor predict individual article citation rate? Aesthetic Surg. J.
Asante, B.O., Osei, M.K., Dankyi, A.A., Berchie, J.N., Mochiah, M.B., Lamptey, J.N.L., Bolfrey-Arku, G., 2013. Producer characteristics and determinants of technical efficiency of tomato-based production systems in Ghana. J. Dev. Agric. Econ. 5, 92–103.
Azoulay, P., Lynn, F.B., 2020. Self-citation, Cumulative Advantage, and Gender Inequality in Science (No. W26893). National Bureau of Economic Research. Working Paper 26893. Available at:http://www.nber.org/papers/w26893. (Accessed 20 March 2020).
Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag. Sci. 30, 1078–1092.
Batooli, Z., Janavi, E., Ravandi, S.N., 2017. The impact of ResearchGate indicators on increasing citation counts of top Clinical Medicine articles in Web of Science: a Comparative study of Iranian and Turkish Researchers. QJ Knowledge Inf Management 3, 83–93.
Bellemare, M.F., Wichman, C.J., 2019. Elasticities and the inverse hyperbolic sine transformation. Oxf. Bull. Econ. Stat.
Belotti, F., Deb, P., Manning, W.G., Norton, E.C., 2015. twopm: two-part models. STATA J. 15, 3–20.
Belsley, D.A., Kuh, E., Welsch, R.E., 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Wiley, New York.
Bornmann, L., 2017. Is collaboration among scientists related to the citation impact of papers because their quality increases with collaboration? An analysis based on data from F1000Prime and normalized citation scores. J. Assoc. Info. Sci. Technol. 68, 1036–1047.
Bornmann, L., Daniel, H.D., 2006. Selecting scientific excellence through committee peer review-A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. Scientometrics 68 (3), 427–440.
Bornmann, L., Schier, H., Marx, W., Daniel, H.D., 2012. What factors determine citation counts of publications in chemistry besides their quality? J.Info. 6, 11–18.
Bornmann, L., Wray, K.B., Haunschild, R., 2020. Citation concept analysis (CCA): a new form of citation analysis revealing the usefulness of concepts for other researchers illustrated by exemplary case studies including classic books by Thomas S. Kuhn and Karl R. Popper. Scientometrics 122, 1051–1074.
Bosquet, C., Combes, P.P., 2013. Are academics who publish more also more cited? Individual determinants of publication and citation records. Scientometrics 97, 831–857.
Bozzo, A., Oitment, C., Evaniew, N., Ghert, M., 2017. The journal impact factor of orthopaedic journals does not predict individual paper citation rate. J. Am. Acad. Orthop. Surg. Glob. Res. Rev. 1 (2), 2–6.
Burbidge, J.B., Magee, L., Robb, A.L., 1988. Alternative transformations to handle extreme values of the dependent variable. J. Am. Stat. Assoc. 83, 123–127.
Carboni, O.A., 2012. An empirical investigation of the determinants of R&D cooperation: an application of the inverse hyperbolic sine transformation. Res. Econ. 66, 131–141.
Card, D., DellaVigna, S., 2017. What do editors maximize? Evidence from four economics journals. Rev. Econ. Stat. 102, 195–217.
Card, D., DellaVigna, S., Funk, P., Iriberri, N., 2020. Are referees and editors in economics gender neutral? Q. J. Econ. 135, 269–327.
Cartwright, V.A., Savino, P.J., 2009. Ophthalmology journals and the ether: considering journal impact factor and citation analysis in context. Clin. Exp. Ophthal. 37 (9), 833–835.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision-making units. Eur. J. Oper. Res. 2, 429–444.
Cherrier, B., 2017. Classifying economics: a history of the JEL codes. J. Econ. Lit. 55, 545–579.
Ciminata, G., Geue, C., Langhorne, P., Wu, O., 2020. A two-part model to estimate inpatient, outpatient, prescribing and care home costs associated with atrial fibrillation in Scotland. BMJ Open 10, e028575.
Constantin, P.D., Martin, D.L., Rivera, R.Y., De, E.B.B., 2009. Cobb-Douglas, translog stochastic production function and data envelopment analysis in total factor productivity in Brazilian agribusiness. J. Operation. Supply Chain Manag. (JOSCM) 2, 20–33.
Cozzens, S.E., 1985. Comparing the sciences: citation context analysis of papers from neuropharmacology and the sociology of science. Soc. Stud. Sci. 15, 127–153.
Cozzens, S.E., 1981. Taking the measure of science: a review of citation theories. Newsletter, Int. Soc. Sociology Knowledge 7, 16–21.
Cozzens, S.E., 1989. What do citations count? The rhetoric-first model. Scientometrics 15 (5–6), 437–447.
Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica: J. Econom. Soc. 39, 829–844.
Cronin, B., 1981. The need for a theory of citing,. J. Doc. 37, 16–24.
Cronin, B., 1984. The Citation Process. The Role and Significance of Citations in Scientific Communication. Taylor Graham, London.
Daraio, C., Kerstens, K., Nepomuceno, T., Sickles, R.C., 2020. Empirical surveys of frontier applications: a meta-review. Int. Trans. Oper. Res. 27, 709–738.
Dean, E.B., French, M.T., Mortensen, K., 2020. Food insecurity, health care utilization, and health care expenditures. Health Serv. Res.
Dehghan, M., Balouchi, H., Yadavi, A., 2020. Improvement of seed quality of wheat (*Triticum aestivum*) as affected by brassinolide under different irrigation regimes. J. Crop Sci. Biotechnol. 23, 137–148.
Djokoto, J.G., Gidiglo, F.K., Srofenyoh, F.Y., Agyei-Henaku, K.A.A., Arthur, A.A.A., Badu-Prah, C., 2020. Sectoral and spatio-temporal differentiation in technical efficiency: a meta-regression. Cogent Econom. Finance 8, 1773659.
Dorta-González, P., Santana-Jiménez, Y., 2018. Prevalence and citation advantage of gold open access in the subject areas of the Scopus database. Res. Eval. 27, 1–15.
Duan, N., Manning, W.G., Morris, C.N., Newhouse, J.P., 1984. Choosing between the sample-selection model and the multi-part model. J. Bus. Econ. Stat. 2, 283–289.
Elkins, M.R., Maher, C.G., Herbert, R.D., Moseley, A.M., Sherrington, C., 2010. Correlation between the journal impact factor and three other journal citation indices. Scientometrics 85, 81–93.
Fare, R., Grosskopf, S., Lovell, C.A.K., 1983. The structure of technical efficiency. Scand. J. Econ. 85, 181–190.
Farewell, V.T., Long, D.L., Tom, B.D.M., Yiu, S., Su, L., 2017. Two-part and related regression models for longitudinal data. Annual Rev. Statistic. Appl. 4, 283–315.
Fleming, L., 2001. Recombinant uncertainty in technological search. Manag. Sci. 47, 117–132.
Frenken, K., Hölzl, W., De Vor, F., 2005. The citation impact of research collaborations: the case of European biotechnology and applied microbiology (1988–2002). J. Eng. Technol. Manag. 22, 9–30.
Frenken, K., Ponds, R., Van Oort, F., 2010. The citation impact of research collaboration in science-based industries: a spatial-institutional analysis. Pap. Reg. Sci. 89, 271–351.
Garfield, E., 1979. Citation Indexing. Its Theory and Application in Science, Technology, and Humanities. John Wiley &. Sons, New York.
Garfield, E., Malin, M., Small, H., 1979. Citation Data as Science Indicators. John Wiley & Sons, New York, pp. 179–207 chapter 8.
Gilbert Nigel, G., 1977. Referencing as persuasion. Soc. Stud. Sci. 7, 113–122.
Gordon, R.A., 1968. Issues in multiple regression. Am. J. Sociol. 73, 592–616.
Greene, W.H., 1993. Econometric Analysis, second ed. Macmillan, New York.
Gusenbauer, M., 2019. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. Scientometrics 118, 177–214.
Gusenbauer, M., Haddaway, N.R., 2020. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of google scholar, PubMed and 26 other resources. Res. Synth. Methods 11, 181–217.
Gutiérrez, E., Lozano, S., Guillén, J., 2020. Efficiency data analysis in EU aquaculture production. Aquaculture 520.
Halvorsen, R., Palmquist, R., 1980. The interpretation of dummy variables in semilogarithmic equations. Am. Econ. Rev. 70, 474–475.
Hassan, N.R., Loebbecke, C., 2017. Engaging scientometrics in information systems. J. Inf. Technol. 32 (1), 85–109.
Hengel, E., Moon, E., 2020. Gender and quality at top economics journals. Working paper No. 202001. https://erinhengel.github.io/Gender-Quality/quality.pdf.
Hubbard, D.E., 2017. Open access citation advantage? A local study at a large research university. Proc. Assoc. Info. Sci. Technol. 54, 712–713.
Kaplan, N., 1965. The norms of citation behaviour: prolegomena to the footnote. Am. Doc. 16, 179–184.
Kaur, A., Rattan, G.K., 2018. Citation analysis of theses in economics submitted to Punjabi university, Patiala during 2000-2014. DESIDOC J. Library Info. Technol. 38, 192.
Kosnik, L.R., 2018. A survey of JEL codes: what do they mean and are they used consistently? J. Econ. Surv. 32, 249–272.

Lampe, H.W., Hilgers, D., 2015. Trajectories of efficiency measurement: a bibliometric analysis of DEA and SFA. Eur. J. Oper. Res. 240, 1–21.
Lartey, S.T., de Graaff, B., Magnussen, C.G., Boateng, G.O., Aikins, M., Minicuci, N., Kowal, C., Si, L., Palmer, A.J., 2020. Health service utilisation and direct healthcare costs associated with obesity in older adult population in Ghana. Health Pol. Plann. 35, 199–209.
Latour, B., 1987. Science in Action: How to Follow Scientists and Engineers through Society. Harvard University Press.
Lavin, C.W., McNab, R.M., Sullivan, R.S., 2017. The long term effects of an aging fleet on operational availability and cost: evidence from the US coast guard. Defence Peace Econ. 28 (6), 634–651.
Lee, J.D., Liu, S.Y., Domeyer, J., DinparastDjadid, A., 2019. Assessing drivers' trust of automated vehicle driving styles with a two-part mixed model of intervention tendency and magnitude. Hum. Factors.
Lee, S.Y., Lee, S., Jun, S.H., 2010. Author and article characteristics, journal quality and citation in economic research. Appl. Econ. Lett. 17, 1697–1701.
Leydesdorff, L., 1998. Theories of citation? Scientometrics 43 (1), 5–25.
Leydesdorff, L., Bornmann, L., Wagner, C.S., 2019. The relative influences of government funding and international collaboration on citation impact. J. Assoc. Info. Sci. Technol. 70, 198–201.
Liu, M., 1993. Progress in documentation the complexities of citation practice: a review of citation studies. J. Doc. 49 (4), 370–408.
Liu, J.S., Lu, L.Y., Lu, W.M., Lin, B.J., 2013. Data envelopment analysis 1978–2010: a citation-based literature survey. Omega 41, 3–15.
Liu, Y., Rousseau, R., 2013. Interestingness and the essence of citation. J. Doc. 69, 580–589.
Luukkonen, T., 1991. Citation indicators and peer review: their time-scales, criteria of evaluation, and biases. Res. Evaluation 1 (1), 21–30.
Luukkonen, T., 1997. Why has Latour's theory of citations been ignored by the bibliometric community? Discussion of sociological interpretations of citation analysis. Scientometrics 38, 27–37.
MacKinnon, J.G., Magee, L., 1990. Transforming the dependent variable in regression models. Int. Econ. Rev. 31, 315–339.
MacRoberts, M.H., MacRoberts, B.R., 1987. Another test of the normative theory of citing. J. Am. Soc. Inf. Sci. 38, 305.
Maz-Machado, A., Jiménez-Fanjul, N., 2018. Collaboration and citation analysis within social sciences: a comparative analysis between two fields. In: Scientometrics. IntechOpen, pp. 65–82.
Meeusen, W., Van Den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. Int. Econ. Rev. 18, 435–444.
Merton, R.K., 1979. Foreword. In: Garfield, E. (Ed.), Citation Indexing: its Theory and Application in Science,.
Mihaylova, B., Briggs, A., O'Hagan, A., Thompson, S.G., 2011. Review of statistical methods for analysing healthcare resources and costs. Health Econ. 20, 897–916.
Mulkay, M.J., 1974. Methodology in the sociology of science: some reflections on the study of radio astronomy. Information 13, 107–119.
Narin, F., Whitlow, E.S., 1990. Measurement of Scientific Cooperation and Co-authorship in CEC-related Areas of Science (Vol. 1). Commission of European Communities [Research Evaluation (EUR 12900 EN)], Luxembourg.
Narin, F., Stevens, K., Whitlow, E.S., 1991. Scientific co-operation in Europe and the citation of multinationally authored papers. Scientometrics 21 (3), 313–323.
Niyazov, Y., Vogel, C., Price, R., Lund, B., Judd, D., Akil, A., Shron, M., 2016. Open access meets discoverability: citations to articles posted to Academia.edu. PloS One 11, e0148257.
Nomaler, Ö., Frenken, K., Heimeriks, G., 2013. Do more distant collaborations have more citation impact? J.Info. 7, 966–971.
Nuti, S.V., Ranasinghe, I., Murugiah, K., Shojaee, A., Li, S.X., Krumholz, H.M., 2015. Association between journal citation distribution and impact factor: a novel application of the Gini coefficient. J. Am. Coll. Cardiol. 65, 1711–1712.
O'brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quantity 41, 673–690.
Orduña-Malea, E., Ayllón, J.M., Martín-Martín, A., López-Cózar, E.D., 2015. Methods for estimating the size of Google Scholar. Scientometrics 104 (3), 931–949.
Pence, K.M., 2006. The role of wealth transformations: an application to estimating the effect of tax incentives on saving. B E J. Econ. Anal. Pol. 5, 1–26.
Peritz, B.C., 1992. On the objectives of citation analysis: problems of theory and method. J. Am. Soc. Inf. Sci. 43, 448–451.
Persson, O., Glänzel, W., Danell, R., 2004. Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. Scientometrics 60, 421–432.
Price, D.D.S., 1976. A general theory of bibliometric and other cumulative advantage processes. J. Am. Soc. Inf. Sci. 27, 292–306.
Price, D. de Solla, 1970. Citation measures of hard science, soft science, technology and nonscience. In: Nelson, C.E., Pollock, D.K. (Eds.), Communication Among Scientists and Engineers. Heath Lexington Books, Lexington, MA, pp. 3–22.
Ramalho, J.J.S., da Silva, J.V., 2009. A two-part fractional regression model for the financial leverage decisions of micro, small, medium and large firms. Quant. Finance 9, 621–636.
Ravetz, J.R., 1971. Scientific Knowledge and its Social Problems. Clarendon Press, Oxford.
Razumova, I.K., Kuznetsov, A., 2019. Impact of open access models on citation metrics. J. Info. Sci. Theory Practice 7, 23–31.
Ronda-Pupo, G.A., Katz, J.S., 2017. The scaling relationship between citation-based performance and coauthorship patterns in natural sciences. J. Assoc. Info. Sci. Technol. 68, 1257–1265.
Rosseau, R., 1998. Citation analysis as a theory of friction or polluted air? Scientometrics 43, 63–67.
Rousseau, R., 1992. Why am I not cited or why are multi-authored papers more cited than others? J. Doc. 48, 79–80.
Ruhinduka, R.D., Alem, Y., Eggert, H., Lybbert, T., 2020. Smallholder rice farmers' post-harvest decisions: preferences and structural factors. Eur. Rev. Agric. Econ.
Sababi, M., Marashi, S.A., Pourmajidian, M., Pourtabatabaei, S.S., Darki, F., Sadrzadeh, M.R., Jamalkhah, M., 2017. How accessibility influences citation counts: the case of citations to the full text articles available from ResearchGate. RT. A J. Res. Pol. Evaluation 5, 1–12.
Shengbo, L., 2015. The theory and method of citation content analysis. Inf. Stud.: Theory Appl. 10.
Small, H.G., 1978. Cited documents as concept symbols. Soc. Stud. Sci. 8, 327–340.
Small, H., 1999. Visualizing science by citation mapping. J. Am. Soc. Inform. Sci. 50 (9), 799–813.
Stanley, T.D., Doucouliagos, H., Giles, M., Heckemeyer, J.H., Johnston, R.J., Laroche, P., Nelson, J.P., Paldam, M., Poot, J., Pugh, G., Rosenberger, R.S., Rost, K., 2013. Meta-analysis of economics research reporting guidelines. J. Econ. Surv. 27, 390–394.
Staudt, J., 2018. Mandating access: assessing the NIH's public access policy. Available at SSRN 3238015.
Struck, D.B., Durning, M., Roberge, G., Campbell, D., 2018, September. Modelling the effects of open access, gender and collaboration on citation outcomes: replicating, expanding and drilling. In: 23rd International Conference on Science and Technology Indicators (STI 2018). Centre for Science and Technology Studies (CWTS), Leiden, The Netherlands. September 12-14, 2018.
Thelwall, M., Maflahi, N., 2019. Academic collaboration rates and citation associations vary substantially between countries and fields. J. Assoc. Info. Sci. Technol. 1–11.
United Nations, 2008. International standard industrial classification of all economic activities revision 4. Statistical Papers, Series M No. 4/Rev.4. United Nations, New York, US.
Van Raan, A.F., 1998. In matters of quantitative studies of science, the fault of theorists is offering too little and asking too much. Scientometrics 43, 129–139.
Waltman, L., 2016. A review of the literature on citation impact indicators. J. Info. 10 (2), 365–391.
Wohlrabe, K., Birkmeier, D., 2014. Do Open Access Articles in Economics Have a Citation Advantage? MPRA Paper No. 56842. Online:https://mpra.ub.uni-muenchen.de/56842/.
Wongkhae, K., Srithongdeang, C., Sittithawan, P., 2017. Regression model of citation counts and reference impact index. Chiang Mai Univ. J. Econom. 21, 105–125.
Wooldridge, J.M., 2009. Introductory Econometrics: A Modern Approach, 4. South-Western.
Wouters, P., 1998. The signs of science. Scientometrics 41, 225–241.
Wulff, J.N., 2019. Generalized two-part fractional regression with cmp. STATA J. 19 (2), 375–389.
Yasin, A., Fatima, R., Wen, L., Afzal, W., Azhar, M., Torkar, R., 2020. On using grey literature and Google Scholar in systematic literature reviews in software engineering. IEEE Access 8.
Yenibehit, N., Murshed, M., Islam, M., 2019. Assessment of technical efficiency of layer production in Mampong municipality: stochastic frontier approach. Curr. Res. Agric. Sci. 6, 20–28.
Yue, W., Wilson, C.S., 2004. Measuring the citation impact of research journals in clinical neurology: a structural equation modelling analysis. Scientometrics 60, 317–332.
Zhao, J., Zhao, Y., Xiang, L., Khanal, V., Binns, C.W., Lee, A.H., 2020. A two-part mixed-effects model for analysing clustered time-to-event data with clumping at zero. Comput. Methods Progr. Biomed. 187.