Impact of academic authorship characteristics on article citations*

Authors: PHILIPP OTTO
– European University Viadrina,
Germany (otto@europa-uni.de)

PHILIPP OTTO
– European University Viadrina,
Germany (potto@europa-uni.de)

Abstract:

• Scientific self-evaluation practices are increasingly built on citation counts. Citation practices for the top journals in economics, psychology, and statistics illustrate article characteristics that influence citation frequencies. Citation counts differ between the investigated disciplines, with economics attracting the most citations and statistics the least. Although articles in statistics are cited less frequently, its proportion of uncited articles is the smallest of all three disciplines. Academic authorship characteristics clearly influence the number of citations. Having authors alphabetically ordered, a practice differently present in the investigated disciplines, increases citations. Further, the more authors there are, the more the article is cited, and a first author with a common surname has positive effects on citation counts, whereas two or more authors sharing a surname attracts fewer citations. In addition, the shorter the article's title, the higher the number of citations.

Key-Words:

• *Scientometrics; Publication Index; Citation Characteristics; Popular Author Names; Alphabetical Authorship.*

AMS Subject Classification:

• 62C25, 91C05.

---

*To the best of our knowledge, this is the first published article where both authors share their surname and given name, while working at the same university. Thus, the two authors are largely indistinguishable, which highlights the importance of individual author identifiers like ORCID.

"If men define situations as real, they are real in their consequences."

(William Isaac Thomas & Dorothy Swaine Thomas 1859, p. 572)

---

1.   INTRODUCTION

Being cited is typically good news for the author(s) of a paper. However, the reference made could be rather critical. In any case, the number of citations reflects the academic impact of an article, and citation counts often provide an initial estimate of the quality of the cited publication, its author(s), and the publishing journal. Because journal rankings and, therefore, academic success are increasingly based on citation counts, the central aim of journal editors appears to be to select articles with the highest citation count expectation (cf., Bornmann et al. 2011). Whereas the practice of quantifying the number of achieved citations in published work is widespread and appears rather useful, citation criteria are manifold and can potentially be self-supporting.

Generally, citation rates are difficult to predict. In this paper, potential drivers are investigated on an exemplary basis for the highest SCImago-ranked journals in economics, psychology, and statistics. Even after ten years, a large proportion (12.4%) of articles were not cited, and half of the articles in the top-ranked journals remained below 20 citations, whereas the total number of citations is slightly above 200 on average. Considering average citations per year, the maximum increase in citations is reached somewhere after 11 years (see Figure 1). This leads to the question of whether there are any identifiable criteria that can explain higher citation counts?

The most common dependency is that the more an article has been cited in the past, the more it will be cited in the future (cf. Stegehuis et al. 2015). Furthermore, a typical article citation curve describes a steady increase over its life cycle. Within approximately three years, an article typically gains momentum (or lack thereof), then reaches a top level of citations somewhere between 10 and 15 years. Thereafter, the majority of articles are cited less frequently.[1] Various factors can be investigated to compare the above-median cited articles against those below. We quantify some easily available article differentials, with a concentration on authorship characteristics, namely *research discipline*, *years since publication*, *title length*, *number of authors*, *alphabetically ordered authors*, *author name-sharing*, and *common author name*). Beginning with a specification of the potential influences and postulating canonical regularities, we provide an empirical analysis using a freely-available data source with an accordingly adapted statistical model and present the results for the investigated dependencies. In the

---

[1]A more general description of citation changes over time, with more profound numbers on passing critical thresholds to develop a momentum, would require time-series data. Investigations that account for other temporal influences, such as citation density or prolonging increases in citations are provided by Quandt (1976) or Parolo et al. (2015).
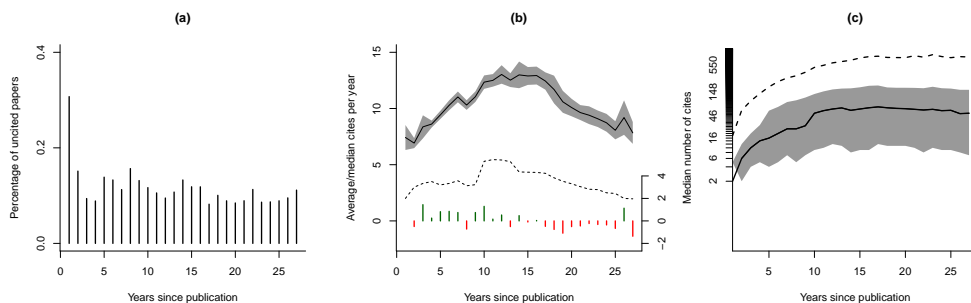
Figure 1: Temporal dynamics of the total number of citations per year since publication. (a) percentage of uncited articles; (b) average citations (solid line, with 95%-confidence intervals as shaded area) and median citations (dashed line) per year depicted for papers, which have been published $1, 2, ..., 27$ years ago, with absolute temporal differences per year as red/green-colored bars; (c) median total number of citations after $1, 2, ..., 27$ years with the shaded area representing the interquartile range and the 95% quantile as a dashed line.

conclusion, the postulated regularities are critically evaluated, how these results relate to other regularities reported in the literature is discussed, and an outlook on the future development of applicable article quality criteria is provided.

2. CITATION CRITERIA AND POSTULATED DEPENDENCIES

The hereby proposed citation criteria introduce alternative measures for explaining citation counts, which are derived historically, structurally, or purely descriptively. All the tested criteria are easily quantifiable and can be divided into the following two categories: structural regularities, or purely authorship-related characteristics. This shifts the focus from quality or relevance toward other criteria as the ones being responsible for citation counts. As an implicit test, it refutes the discussion on the usefulness of derived empirical indicators for academic success, such as the Hirsch (2005) index and others (compare for example Lindsey 1989), but also illustrates potential regularities as to the ways researchers are citing each other's work.

2.1. Structural regularities

Differences in academic disciplines provide a starting point in the evaluation of article characteristics to find regularities in citing practices. Here, economic, psychology, and statistics publications were used to study discipline-specific differences, as well as broader influences on citation frequencies.

The following exemplary regularities were provided ad hoc: psychological publications would be cited more often (mainly in other disciplines) due to a generally larger public interest in their research topic and strong interdisciplinary focus (compare interdisciplinary citations in Jacobs 2013). Statistics is the smallest discipline and, therefore, citations were expected to be less frequent, although statistics are used for empirical analyses in all disciplines. This postulates a regularity that can be summarized as

Hypothesis 1.   Citation frequencies vary over research disciplines with (a) being higher for psychology publications, and (b) lower for statistics publications.

Other characteristics can be article specific and illustrate a direct structural dependency with citation frequencies. Two discipline-independent influences were proposed with opposing regularities: citation frequencies increase with the *years since publication* and decrease with the *title length of the article*. Naturally, it takes time for articles to be cited and for the academic community to acknowledge new work. However, one could also expect a slowdown several years after the time of publication, due to decreased novelty. Another issue that was included is simplicity. An anticipated effect is based on information processing and recall. The *title length of the article* serves as an indicator to investigate this kind of influence. Bounded rationality, in the form of limitations when recalling more complex article titles, could lead to lower citation counts. These two apparent article characteristics needed to be controlled, in addition to the differences between the research disciplines, when investigating the following influences.

2.2. Authorship characteristics

Authorship characteristics might also affect citation frequencies. These characteristics could result from academic practices or other easily identifiable article differentials. Thus, the guiding question was, how much variance in citation frequencies can be explained by extrinsic article characteristics related to authorship. This would be in addition to structural influences and the article's quality as the fundamental value.

The first source for identifiable article differentials is academic differences

based on the cultural and historical development of respective research disciplines. A prominent example in this regard would be how authors are ordered in a joint publication. Some disciplines prefer purely alphabetical order, whereas others strictly list the author names in the order of the contributed amounts of work. This difference in approach for author listing is exploited by Van Praag and van Praag (2008) and Einav and Yariv (2006), who postulate a positive correlation between the surname initials and the scientific success of the author. The influence of the initial letter of the first author can, thus, be seen as a random characteristic independent of the article's quality.

Our three investigated research fields differ with regard to author listing order. Author listings could be either alphabetical or organized by their respective shares of work (i.e., the first author would be the main author of the article). However, it is not always feasible to distinguish between these two kinds of author listings. A non-alphabetically sorted list of authors does not automatically imply that the first author contributed the most, and in an alphabetically sorted list of authors, the first author could still be the main contributing author. For simplification, Figure 2 illustrates this relation for articles with two authors. Plot (a) shows the percentage of articles in which the authors are listed alphabetically. Van Praag and van Praag (2008) computed the probability of an alphabetical ordering for uniformly distributed first letters. However, the chance of having a surname with the initial letter being 'A' differs from that of having the initial letter 'Z'. Hence, in our data set, we used the observed frequencies of the first letters of all surnames as a proxy for the natural distribution of initial letters. The ratio between the observed percentages of alphabetically ordered authors, and this baseline probability can be seen as the percentage of authors intentionally sorted by the first letter of their surnames. This further implies that the authors of the remaining articles are listed in a non-alphabetical way – potentially to reflect the amount of contributed work. The accordingly estimated proportions of intentionally alphabetically ordered authors are shown in Figure 2 (b), which were strictly lower in psychology when compared with economics and statistics. One can conclude that the first author is most likely to be the main author for articles published in the top psychology journals, whereas in economics and statistics, both authorship orderings coexist.[2] Note that only the first letter of the surname is compared. Names with the same first letter are considered as being alphabetically ordered, although this includes the curiosity that, if all authors have the same surname, they are considered as being alphabetically ordered, although these are at the same time non-alphabetically ordered.

In addition to the citation differences between the three investigated disciplines, publication practices could affect an article's citation count. The two

---

[2]As the estimated frequencies from our data set could be biased, Appendix A provides a comparison of these results to the distribution of UK surnames, as reported by Gray (1958), and for the top 100 surnames in the United States of America (provided by the U.S. Census Bureaus for the year 2000), thereby confirming these regularities.
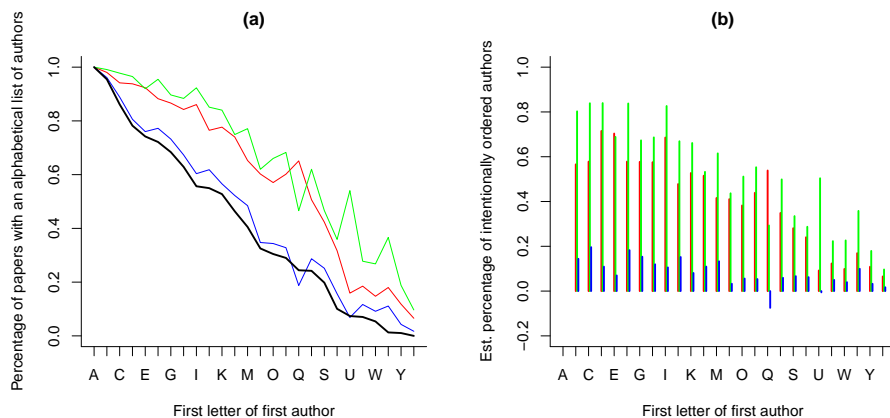
Figure 2:
Percentage of articles with two authors having alphabetically ordered names separated by the initial letter of the first author. (a) shows the percentage of ordered lists of authors in economics (red), statistics (green), and psychology (blue). The bold line depicts the probability of two random surnames being in alphabetical order. (b) ratio between observed frequencies and the expected base probability (black baseline) illustrates the proportion of intentionally alphabetically ordered authors.

different ways of ordering authors might directly influence its number of citations because the main author is not easily identifiable with *alphabetically ordered authors*, and the allocation of the main work to one specific versus various researchers might influence its citation.

Hypothesis 2.   Citation frequencies change when the main author is listed as the first author of the article.

The relation between citation counts and surname familiarity is included in the analysis as another test for the influence of recall simplicity. The top 100 U.S. surnames served as a proxy for *common author names*.[3]

Hypothesis 3.   Citation frequencies increase with the first author having a common surname.

Another simplicity-related claim goes back to Goodman et al. (2015), who investigated a descriptive curiosity of authors sharing surnames. Sources for name doubling, or more generally *author name-sharing*, could be for various reasons and could also directly link to citation counts. Without knowing why the same

---

[3]This list also includes popular surnames from other nationalities (e.g., Lee, Nguyen, or Rodriquez). In addition, we considered the soundex of all names to account for different spellings such as Li, Lee, or Liu, but this opposes a unique author identification and, thereby, the postulate of recall simplicity.

name occurs twice (or even more often), we argue that these articles are easier to remember and to recall.

Hypothesis 4. Citation frequencies increase when authors share their surnames.

A more universal relationship is hypothesized for authorship with regard to the number of people involved with the published research. The *number of authors* is expected to show a direct relationship with citation counts.

Hypothesis 5. Citation frequencies increase with the number of listed authors for an article.

With more authors, the new information spreads faster and can be expected to be better connected within the respective scientific communities – not to mention direct (or reciprocal) self-citations.

3. EMPIRICAL DATA ANALYSIS

The systematic rating of evoked citations increasingly influences the scientific evaluation process, ranging from the rankings of individual publications to that of authors and journals. A practical advantage is that citations can easily be retrieved, in addition to diverse article characteristics.[4] The predictive variables of interest are the *research discipline*, *years since publication*, *title length*, *number of authors*, *alphabetically ordered authors*, *author name-sharing*, and *common author name*.

3.1. Data and Descriptive Statistics

The data analysis was based on $196,365$ journal articles that were published in 115 journals from 1990 to 2016. For each, we observed the current citation count as well as various article characteristics. To be precise, the focus was on the highest-ranked journals in three scientific fields, namely economics, psychology, and statistics. The definition of journals belonging to the top journals, to be included in the following analysis, is based on the SCImago journal ranking within the respective subject areas:

- "Economics, Econometrics and Finance:" top ten journals of each subcategory (except "Science" as not being a mainly economic journal)

---

[4]Different elicitation methods are described more broadly in Ball (2014).

- "Psychology:" top ten journals of each subcategory

- "Statistics, Probability and Uncertainty:" top quartile journals (as already a "subcategory").

All included journals are listed in Table 1 (31 from economics, 57 from psychology, and 27 from statistics), with the number of articles, the average SCImago journal ranking index (SJR), the average Hirsch index (H), and the average citations per document for each of the three investigated research areas. The number of total citations recorded until November 2017 serves as a performance measure of each article. To be more specific, citation counts reported by Microsoft Academic Search (MAS) are used as the dependent variable. These counts partly incorporate statistical models based on network data to provide more accurate citation counts; a more detailed discussion of the data set and the MAS citation count is provided in Appendix A.

For the empirical analysis of the postulated hypotheses, we use the current citation counts of all papers published within these journals and the above-mentioned time period. Hence, the citation counts are cumulated values for each individual paper, but independent across time because each paper appears only once in the sample. Figure 1 (a) depicts the percentage of uncited articles with respect to the elapsed *years since publication* (in full years). This ratio decreases from thirty percent for all publications in the year of publication (i.e., 2016) to approximately twelve percent within the first three years. The proportion of articles not cited remains stable thereafter, whereas the total number of citations increases over time. The positive growth rate lasts for about 11 years after publication.

The annual average and median citations depicted in plot (b) of Figure 1 have their peaks after 11 years, which implies declining growth rates afterward. However, it is important to note that we have independent samples over time, such that the downslope is partly due to the generally increasing number of citations. For comparison, we also depict the lower quartiles, medians, upper quartiles, and 95%-quantiles of the total citation counts over the elapsed time since publication on a log-scale in Figure 1 (c). This supports the assumption that the number of new citations increases in the beginning but reduces with decreasing novelty, and the latter effect seems to be strengthened by an overall increase in the number of citations over the years since 1990 (i.e., older articles are cited less often over their citation life-span). Moreover, Table 2 summarizes the descriptive statistics for the central variables of the regression: the number of citations, percentage of uncited articles, average years since publication, and number of authors (36.7% with one author and 25.7% with two authors). In addition, the average title length is included as the number of characters in the title of the article. Author name-sharing occurred in 0.2% of all included articles.

3.2. Model

Because more than ten percent of the articles were not cited within the investigated time frame, the statistical model needs to account for this excess of non-citations. For our data, a zero-inflated negative binomial model was used because it provided a comparatively better fit than other models (e.g., a zero-inflated Poisson model), which is further supported by the Ord plot (see Ord 1967). Please see Appendix B for a more detailed discussion of this distributional choice.

To define the statistical model, we introduce a random variable Y for the citation counts. The observations of Y are denoted by y. Then, the conditional probability of Y is given by

$$P(Y = y | \mathbf{X}_z, \mathbf{X}_c, \boldsymbol{\beta}_z, \boldsymbol{\beta}_c) =$$

(3.1) $\qquad P_z(Y = 0 | \mathbf{X}_z, \boldsymbol{\beta}_z) I_{\{0\}}(y) + (1 - P_z(0 | \mathbf{X}_z, \boldsymbol{\beta}_z)) P_c(Y = y | \mathbf{X}_c, \boldsymbol{\beta}_c),$

where $\mathbf{X}_z$ and $\mathbf{X}_p$ are the matrices of explanatory variables for the probability of $Y = 0$ (index z) and $Y = y \geq 0$ (index c). The respective coefficients for these regressors are $\boldsymbol{\beta}_c$ and $\boldsymbol{\beta}_z$. Moreover, $I_A(x)$ stands for the indicator function on a set A. Whereas P_z describes the conditional probability for $Y = 0$, the probability density of P_c defines the number of citations. For our analysis, we assume that P_c is a negative binomial distribution, i.e.,

$$P_c(Y = y | \mathbf{X}_c, \boldsymbol{\beta}_c) = \frac{\Gamma(\theta + y)}{\Gamma(y + 1)\Gamma(\theta)} r^y (1 - r)^\theta \qquad \text{with} \qquad r = \frac{\exp(\mathbf{X}_c \boldsymbol{\beta}_c)}{\exp(\mathbf{X}_c \boldsymbol{\beta}_c) + \theta}.$$

Due to the methodological separation of articles into cited and uncited, it is possible to distinguish two different effects: the predictive variable $\mathbf{X}_z$, influencing the fact of an article being cited at all, and $\mathbf{X}_c$, influencing the number of citations of a particular work. Corresponding regression coefficients are obtained as maximum-likelihood estimators of a generalized linear model, which is computationally implemented as in Zeileis et al. (2008). The starting values of the iterative maximization of the likelihood function have been chosen by an expectation maximization algorithm.

3.3. Results

All articles were searched for characteristics that explained, firstly, if it was cited at all and, secondly, the number of citations reached.[5] Table 3 shows the

---

[5]Articles with total citations that were above the 95% quantile are neglected to avoid anomalies due to outlying observations.

results of the zero-inflated negative binomial model with parameters estimated by the maximum-likelihood approach (cf. Greene 2003; Zeileis et al. 2008). For this, we included all variables introduced in Section 2 that have a potential influence on citation counts. For simplicity of interpretation of the results, we omit potential interactions between the regressors, which are reported in Appendix C. To allow for a more intuitive interpretation of the regression coefficients, we report the corresponding odds ratios r_i for the count and zero component of the model. These ratios depict the factor by which the expected citation count or probability of being cited changes if the corresponding dummy variable is present or the independent variable is increased by one unit (see Table 3).

3.3.1. Structural regularities

Citation existence and level are highly influenced by the amount of time passed since an article has been published. The older the publication, the higher the likelihood that the publication does not belong to the class of not cited articles, while its citation count is expected to be higher. Thus, *years since publication* increase the likelihood of being cited (negative zero-inflation coefficient $\hat{\beta}_3^z$), as well as the number of citations (positive count coefficient $\hat{\beta}_3^c$). Further, the expected regularities for *title length* are fully confirmed. The longer the title, the more likely it belongs to the uncited articles category and the lower the citation counts. These strong and clear intrinsic influences fully confirm the first two expected regularities, that citation frequencies are indeed determined by the *years since publication* as well as by its *title length*.

Mixed results are observed concerning the differences in the three research disciplines because partly opposite patterns were noted. For Statistics, both coefficients $\hat{\beta}_2^z$ and $\hat{\beta}_2^c$ are negative, which indicates opposite effects. Whereas Statistics has fewer uncited articles when compared with Economics, these articles gather fewer citations. Examining the count model, we see that citation counts were lower in both Psychology (contradicting Hypothesis 1a) and Statistics (supporting Hypothesis 1b). Consequently, Economics attracted the most citations compared to the two other disciplines. Given that an article is cited, Statistics articles were cited less frequently when compared to Economics and Psychology. This fully supports Hypothesis 1b because the respective coefficients of the count model confirm this order, i.e., $0 > \hat{\beta}_1^c > \hat{\beta}_2^c$. Articles in Statistics were cited less often than articles in Psychology ($p < 0.0001$) and articles in Economics ($p < 0.0001$). Moreover, citations in Psychology were lower than in Economics ($p < 0.0001$). These pairwise relations are also supported by Mann-Whitney-U tests on all cited articles (citations > 0). Thus, the postulated order of the disciplines concerning citation frequencies when being cited is confirmed only when comparing Statistics with Psychology or Economics, but not when comparing

Psychology with Economics. The research discipline has a strong influence on the number of citations, but the relations postulated under Hypothesis 1 are only partially confirmed.

3.3.2. Authorship characteristics

Authorship characteristics generally remain influential for citation frequencies, when controlling for structural regularities. However, the empirical findings were not always as hypothesized. Articles having *alphabetically ordered authors* show an opposing effect; these are more inflated by uncited articles, but they are cited more often (i.e., $\hat{\beta}_5^z$ and $\hat{\beta}_5^c$ are positive). Hypothesis 2 is only partially supported. Having the first author as the main author is more likely to attract at least one citation, but this effect is insignificant. Articles where the main author appears as the first author are, in fact, cited significantly less than articles with purely alphabetical ordering.[6]

In contrast, Hypothesis 3 is fully supported. Having a *common author name*, as a first author surname characteristic consistently related to citation likelihood and frequency. Having a common surname increases the probability of being cited. Important here is that judging whether the surname is a common name based on the exact spelling, rather than on its soundex, leads to a better model fit. Thus, the unique spelling of the name seems to be crucial for its recall simplicity. Another unexpected result was observed regarding the influence of *author name-sharing*. For both cases of being cited and the frequency of citations, the relation is in the opposite direction than postulated under Hypothesis 4. Articles that have (for some authors) the same surnames were significantly less likely to be cited, and in cases where they were cited, they are cited significantly less often. Hence, our hypotheses concerning authorship simplicity are only partly confirmed: having a common name has a positive effect, but when authors share the same surname, this is negatively related to citation frequencies. Note that authors randomly sharing a surname is more frequent for popular names.

The strongest influence on citations was the *number of authors*, which increases the likelihood of being cited as well as the number of citations. The negative zero-inflation coefficient ($\hat{\beta}_3^z$) and the positive count coefficient ($\hat{\beta}_3^c$) clearly support Hypothesis 5.

---

[6]Although this effect of alphabetically ordered authors is largely reduced in Psychology, it still has a positive influence across all the considered research disciplines. Interactions with research discipline and their cultural differences in sorting authors is further discussed in Appendix C.

4.   DISCUSSION AND CONCLUSION

Influences on citation counts has received little attention besides noting its fundamental and growing importance for evaluating scientific productivity. Everyday practice simply assumes a direct relation between the gained citations and the importance of the research. This does neglect alternative influences on citation counts. In this regard, various authorship characteristics were evaluated for three research disciplines in social sciences. Without claiming any kind of prominence, systematic regularities can be observed in the data. The *time since publication* is possibly the most important structural component, for which a monotonic increasing relationship is confirmed. To determine an article's citation life (possibly with a critical growth period), however, time series of the citation counts of each article would be required. Although it naturally takes time to acknowledge quality, the duration or speed of this process remains uncertain. Broader issues, such as an overall increase in publications and citations, further complicate this analysis. In addition, fashionable trends are difficult to isolate, particularly in cases where quality intertwines with the novelty of the research topic (compare Van Dalen and Henkens 2001; Webster et al. 2009; Chen 2012). Our empirical results show that the *title length* decreases the likelihood and frequency of being cited. Simplicity might help recognition. A positive relation between an article having a short title and citation counts has already been claimed for economic articles (Bramoullé and Ductor 2018; Gnewuch and Wohlrabe 2017). These results are confirmed here, whereas recognition not only decreases the chance of belonging to the class of uncited articles, but it also increases the number of attracted citations. However, simplicity and recall probability can oppose uniqueness, which might play a role as well. Naturally, the predictive power of such content-free characteristics needs to be investigated in more detail to be applicable because, for example Didegah and Thelwall (2013) claim in a broader study of research disciplines that the length of the title has no significant influence on citation counts.

Differences between the *field of research* (Hypothesis 1) illustrate a more specific regularity in citation frequencies. This potentially originates from other sources than research quality. These differences could have historical reasons or be confounded with the other expected regularities as well as authorship characteristics. We compared articles in Psychology, Economics, and Statistics, where the popularity was expected to decrease in this order (also due to the size of the (sub-)discipline in the case of Statistics). The postulated relationship is not fully reflected in the citation count data. Articles published in the top journals in Psychology are less frequently cited than those in Economics, but publications in Statistics were cited the least. Interestingly, our regression analysis provides a more profound picture. Articles in Statistics are cited less often, but there were also fewer nil citations. These seemingly opposing effects might be due to a flatter distribution pattern, which might also be responsible for the advantage of Economics over Psychology. It is worth noting that only the top journals of

each subject are included in the analysis. A broader sample, of course, might reveal different relations. The proportion of uncited articles can be expected to be more profound and the concentration of citations on fewer articles (such as those in top journals) to be more pronounced in Economics. This is because Economics is more concentrated on a smaller number of leading publications along with a higher impact factor of the top economics journals. This tendency toward the top journals seems to be prolonged (Card and DellaVigna 2013; Heckman and Moktan 2018). Fourcade et al. (2015) claim that Economics is generally more hierarchically organized. Why the pattern of citation counts in Statistics shows a flatter distribution requires further investigation, possibly in comparison to a larger and more diverse number of research fields. In general, explanations for the variety in citation counts has to be searched and accounted for as has been stressed by Varin et al. (2016) regarding cross-citations among highly ranked statistics journals or by Aksnes (2006) for subfields of research in Norway. Radicchi et al. (2008) and Albarrán et al. (2011) provide first approaches to correct citation count evaluations with respect to the field of research.

A central idea put forward here is to isolate various authorship characteristics that can explain part of the observed variation in citations. This could not only lead to a better understanding of the relationship between quality and being cited but also illustrates the potential pitfalls of not being cited. Not all of the included characteristics have a strong effect, and the results do sometimes point in the opposing direction. If articles have *alphabetically ordered authors* (Hypothesis 2), this actually increased the number of citations but reduced the likelihood of being cited at all. This kind of academic tradition, which is more prominent in Economics and Statistics, could represent things other than quality (dominance, conservatism, etc.). Although indirect and only in terms of citation frequencies, this confirms the claim made by Van Praag and van Praag (2008) that authors with names toward the beginning of the alphabet tend to be more successful (under the assumption that an author's future citations directly depend on previous citations).

Author names can also have an influence in terms of their popularity, especially under the expectation of recognition simplicity (Hypothesis 3); namely, that the first author having a *common author name* increases the number of citations, an occurrence that is confirmed by the data. Note that this expectation equally applies to how having an uncommon name (below the 100 most common names benchmark) leads to fewer citations, possibly because it is more difficult to recall unpopular names. Other demographic or personal author characteristics might help to further elaborate upon this kind of relationship. Naturally, author influences that are not investigated here, such as reputation (as for example *author eminence* as in Haslam et al. 2008) or connectivity (Haslam et al. 2008; Vieira and Gomes 2010; Bornmann et al. 2012; Chen 2012; Didegah and Thelwall 2013, as for example *number of references* as in), could play a central role for citation counts. Along the lines of research embedding, the strongest authorship influence on citation counts is the *number of authors* (Hypothesis 5). This is not only the result of self-citations, which have not been distinguished here;

rather, it is attributed to the fact that the more authors there are, the better the interconnectivity and the higher the potential of the paper to be discovered. Thus, the research output is better represented in the respective scientific community, and connections to neighboring fields become more likely. Systematic self- or cross-citations can clearly oppose quality concerns, but dependencies are manifold. For example, collocation effects in the citation networks of authors and institutions can be observed (see Yan and Ding 2012). Still, a larger number of authors can positively affect the quality of an article, due to increased awareness or a more sophisticated cross-checking, for example, but negative effects of co-authorship can also result from this self-selection process (cf. Ductor 2015). Also note that for natural sciences, Onodera and Yoshikane (2015) report only a weak and Bornmann et al. (2012) a negative effect of the number of authors on citation counts. In summary, a better understanding of the different effect strengths of the investigated authorship characteristics is required to be more conclusive here.

Initially most surprising for us was that *author name-sharing* appears to have the opposite effect than expected (Hypothesis 4) because it negatively influences citation counts. Authorship recognition does not appear to be the driving influence. Possibly, this influence of recognizing an article is largely covered by the popularity of the first author's surname because more frequent names already result more often in coauthors sharing their surnames. Further, the list of reasons for authors sharing the name (given by Goodman et al. 2015) provides a plausible answer here. The sources for people having the same name and publishing an article together (i.e., marriage or other family relations) might reduce the quality of its content. However, name-sharing could also be fully coincidental (as in the case of the "Goodmen"). Furthermore, name-sharing might represent narrowness, and internationality has been reported as a factor strongly increasing citations. Documented positive influences are international collaboration (Didegah and Thelwall 2013), authors not sharing the same department (Vieira and Gomes 2010), as well as the article being published in English (Van Dalen and Henkens 2001; Bornmann et al. 2012). This further illustrates the need for systematically distinguishing behavioral influences from those that represent and acknowledge the quality of an article.

Citation indices have been proposed as a heuristic method for informing decision-making on various levels (see for example Perry and Reny 2016; Hamermesh 2018). With diverse drivers influencing citation frequencies, these must be treated even more cautiously. Little has been done to better understand citation behavior, despite it being increasingly crucial in determining academic success. Although it is reasonable to argue that all the articles included in our analysis are of substantial quality because they are published in the top journals of their respective research field, a large proportion are still rarely or not cited at all, whereas other articles strongly pull citations. If specific authorship characteristics are influencing this process, and various data sources exist to evaluate the dependencies here, then these can easily be detected and controlled to better inform decisions. Complementary proxies for research quality are, thus, required to supplement citation indices and journal ranks, both of which are currently solely

based on citation count data.

References

Aksnes, D. W. (2006). Citation rates and perceptions of scientific contribution. *Journal of the American Society for Information Science and Technology*, 57(2):169–185.

Albarrán, P., Crespo, J. A., Ortuño, I., and Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*, 88(2):385–397.

Ball, R. (2014). *Bibliometrie: Einfach-verständlich-nachvollziehbar*. Walter de Gruyter.

Bornmann, L., Mutz, R., Marx, W., Schier, H., and Daniel, H.-D. (2011). A multilevel modelling approach to investigating the predictive validity of editorial decisions: do the editors of a high profile journal select manuscripts that are highly cited after publication? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(4):857–879.

Bornmann, L., Schier, H., Marx, W., and Daniel, H.-D. (2012). What factors determine citation counts of publications in chemistry besides their quality? *Journal of Informetrics*, 6(1):11–18.

Bramoullé, Y. and Ductor, L. (2018). Title length. *Journal of Economic Behavior & Organization*, 150:311–324.

Card, D. and DellaVigna, S. (2013). Nine facts about top journals in economics. *Journal of Economic Literature*, 51(1):144–61.

Chen, C. (2012). Predictive effects of structural variation on citation counts. *Journal of the American Society for Information Science and Technology*, 63(3):431–449.

Didegah, F. and Thelwall, M. (2013). Which factors help authors produce the highest impact research? collaboration, journal and document properties. *Journal of Informetrics*, 7(4):861–873.

Ductor, L. (2015). Does co-authorship lead to higher academic productivity? *Oxford Bulletin of Economics and Statistics*, 77(3):385–407.

Einav, L. and Yariv, L. (2006). What's in a surname? The effects of surname initials on academic success. *Journal of Economic Perspectives*, 20(1):175–187.

Fourcade, M., Ollion, E., and Algan, Y. (2015). The superiority of economists. *Journal of Economic Perspectives*, 29(1):89–114.

Gnewuch, M. and Wohlrabe, K. (2017). characteristics and citations in economics. *Scientometrics*, 110(3):1573–1578.

Goodman, A. C., Goodman, J., Goodman, L., and Goodman, S. (2015). A few Goodmen: Surname-sharing economist coauthors. *Economic Inquiry*, 53(2):1392–1395.

Gray, P. G. (1958). Initial letters of surnames. *Applied Statistics*, 7:58–59.

Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Hamermesh, D. S. (2018). Citations in economics: Measurement, uses, and impacts. *Journal of Economic Literature*, 56(1):115–56.

Haslam, N., Ban, L., Kaufmann, L., Loughnan, S., Peters, K., Whelan, J., and Wilson, S. (2008). What makes an article influential? predicting impact in social and personality psychology. *Scientometrics*, 76(1):169–185.

Heckman, J. J. and Moktan, S. (2018). Publishing and promotion in economics: the tyranny of the top five. Technical report, National Bureau of Economic Research.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46):16569–16572.

Jacobs, J. A. (2013). *In defense of disciplines: Interdisciplinarity and specialization in the research university*. University of Chicago Press.

Lee, Y.-G., Lee, J.-D., Song, Y.-I., and Lee, S.-J. (2007). An in-depth empirical analysis of patent citation counts using zero-inflated count data model: The case of KIST. *Scientometrics*, 70(1):27–39.

Lindsey, D. (1989). Using citation counts as a measure of quality in science measuring what's measurable rather than what's valid. *Scientometrics*, 15(3-4):189–203.

Onodera, N. and Yoshikane, F. (2015). Factors affecting citation rates of research articles. *Journal of the Association for Information Science and Technology*, 66(4):739–764.

Ord, J. (1967). Graphical methods for a class of discrete distributions. *Journal of the Royal Statistical Society. Series A (General)*, pages 232–238.

Parolo, P. D. B., Pan, R. K., Ghosh, R., Huberman, B. A., Kaski, K., and Fortunato, S. (2015). Attention decay in science. *Journal of Informetrics*, 9(4):734–745.

Perry, M. and Reny, P. J. (2016). How to count citations if you must. *American Economic Review*, 106(9):2722–41.

Quandt, R. E. (1976). Some quantitative aspects of the economics journal literature. *Journal of Political Economy*, 84(4, Part 1):741–755.

Radicchi, F., Fortunato, S., and Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences*, 105(45):17268–17272.

Stegehuis, C., Litvak, N., and Waltman, L. (2015). Predicting the long-term citation impact of recent publications. *Journal of Informetrics*, 9(3):642–657.

Thomas, W. I. and Thomas, D. S. (1928). The methodology of behavior study. In Knopf, A. A., editor, *The Child in America: Behavior Problems and Programs*, chapter 13, pages 553–576. online: https://brocku.ca/MeadProject/Thomas/Thomas_1928_13.html, New York.

Van Dalen, H. and Henkens, K. (2001). What makes a scientific article influential? the case of demographers. *Scientometrics*, 50(3):455–482.

Van Praag, C. M. and van Praag, B. (2008). The benefits of being economics professor A (rather than Z). *Economica*, 75(300):782–796.

Varin, C., Cattelan, M., and Firth, D. (2016). Statistical modelling of citation exchange between statistics journals. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(1):1–63.

Vieira, E. S. and Gomes, J. A. (2010). Citations to scientific articles: Its distribution and dependence on the article features. *Journal of Informetrics*, 4(1):1–13.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333.

Wainer, H. (1974). The suspended rootogram and other visual displays: An empirical validation. *The American Statistician*, 28(4):143–145.

Webster, G. D., Jonason, P. K., and Schember, T. O. (2009). Hot topics and popular papers in evolutionary psychology: Analyses of title words and citation counts in evolution and human behavior, 1979–2008. *Evolutionary Psychology*, 7(3):147470490900700301.

Yan, E. and Ding, Y. (2012). Scholarly network similarities: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to each other. *Journal of the American Society for Information Science and Technology*, 63(7):1313–1326.

Zeileis, A., Kleiber, C., and Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, 27(8):1–25.
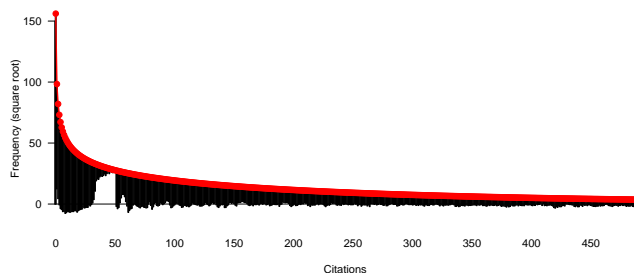
Figure 3:
Rootogram (hanging histogram bars) and best fitting negative binomial distribution colored in red. The gap between 33 and 50 citations is due to the specific reporting of the Microsoft Academic Search program. The total number of citations is shown on a square-root scale.

APPENDICES

A.   DATA SOURCE

Figure 3 visualizes the distribution of the observed counts by a so-called rootogram, depicting the histogram bars pinned to the best-fitting density curve. In this case, we plot the counts against a negative binomial distribution. This figure shows two major issues that need to be addressed. First, uncited articles are excessive because articles cited between one and three times are less frequently observed than expected by a negative binomial distribution. Consequently, we observed such an excess of zero citations that small counts were overestimated. Second, there is a substantial gap in articles for the area between 33 and 50 citation counts. This lack is due to the specific counting approach of Microsoft Academic Search. In particular, the software uses a statistical model based on citation graphs to estimate citation counts, from which the accuracy is lower for all publications just below 50 citations (confirmed by Microsoft Academic Search). Thus, they reported the true citation count only for the remaining publications, for which the predicted count is less than 50. The resulting anomalous pattern for articles cited between 33 and 50 times is rather unsatisfactory. However, the observed effects should not substantially differ, with the main influence on goodness-of-fit measures being based on residuals.

Section 2.2 includes the likelihood of two authors in an article being in alphabetical order, to estimate the proportion of intentionally ordered author lists. The reasons for this calculation would be the observed empirical frequencies of the initial letters, thus resulting in the included articles of the top journals of Economics, Statistics, and Psychology. However, this could be a biased proxy for
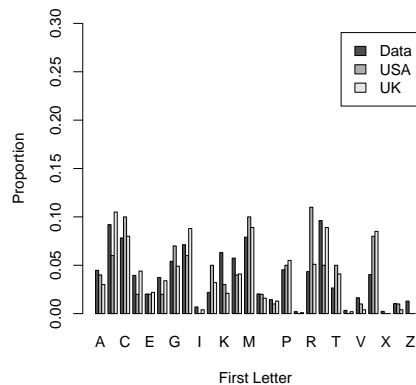
Figure 4: Empirical distribution of the first letter of surnames for our data
set (dark-gray), the top 100 U.S. surnames (gray), and the UK
surnames by Gray (1958) (light-gray).

the true distribution of the first letters of surnames. Hence, we compared these
frequencies to the frequency table published by Gray (1958). In contrast, Gray
(1958) reports the distribution for UK surnames only, which might differ from the
frequency distribution of first letters of surnames globally. To further justify the
results, we also compared our estimated distributions from the data against the
top 100 U.S. surnames from the census in 2002. Figure 4 depicts these empirical
distributions. There are no large differences between the estimated probabilities,
aside for some letters (e.g., 'R' or 'W') where we observe fewer authors in our
data than one would expect when looking at the top 100 U.S. surnames or the
results of Gray (1958). However, this did not affect the main findings. Differences
in the resulting ratios are small, as shown in Figure 5 (analogously to Figure 2),
based on the empirical distribution of UK surnames (also not different for the
100 U.S. surnames census data).

B.   MODEL SELECTION

First impressions of the underlying discrete probability of the citation
counts can be obtained by the so-called Ord's plot (cf. Ord 1967). For our data,
the plot indicates that the data are generated by a negative binomial distribution,
which is also supported by the histogram, or rootogram (e.g. Wainer 1974). Lee
et al. (2007) observed similar behavior for patent citation counts. Comparing a
zero-inflated negative binomial and zero-inflated Poisson model by a Vuong test
(cf. Vuong 1989), the negative binomial model is significantly preferred, with a
test statistic of $|z| = 308.4410$ (uncorrected). Less-complex models, such as the
negative binomial model without zero inflation, can be ruled out due to their
larger information criteria (the Akaike information criterion (AIC) is 1,787,653
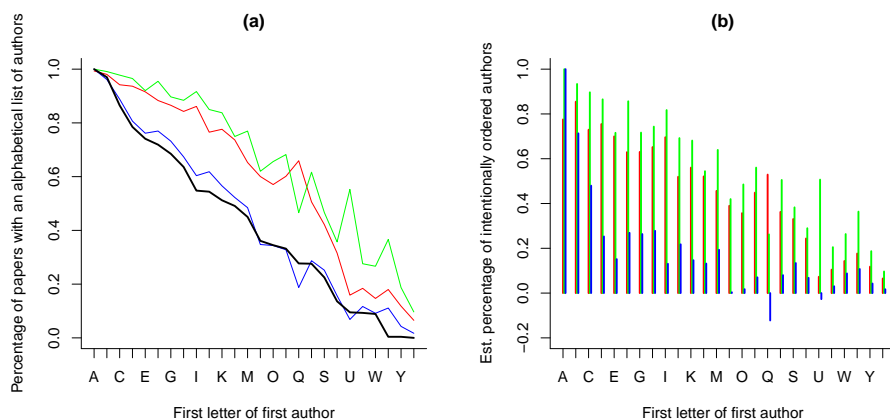
Figure 5: In contrast to Figure 2, we chose the empirical distribution of UK surnames reported by Gray (1958) as a benchmark, i.e., the bold line first plat (a) depicts the probability for two random surnames being in alphabetical order according to this empirical distribution. In the second plot (b), we computed the ratio between the observed frequencies of ordered authors and the estimated probability (black baseline) as an estimate for the percentage of articles that were being intentionally set in alphabetical order.

for the negative binomial model and 1,771,146 for the zero-inflated model).

Moreover, the zero-inflated model allows for the comparison of the probability for being cited and the citation counts across the fields, whereas count data models without zero inflation measure the overall effect. For instance, the estimated coefficients for the indicator variables of research field are -0.235 ($\hat{\beta}_5^c$, Psychology) and -1.055 ($\hat{\beta}_6^c$, Statistics) for a negative binomial, without modeling the inflation of uncited articles. This confirms our results, namely that articles in Psychology are more often cited than in Statistics and that the latter articles are cited the least (in this particular group of the three research disciplines). However, it does not allow for interpretations regarding the excess of uncited articles.

Furthermore, the reported model results (in Table 3 of Section 3.3) include all introduced potential characteristics from Section 2 influencing citation counts as main effects. To provide a model with the best data fit, we also selected covariates and their interactions by stepwise minimizing AIC. The resulting model is discussed next as "model extensions" (in Appendix C).

C. MODEL EXTENSIONS

All results were obtained by a simple regression model, which meant an easier interpretation because we only focused on the direction of the main effects, despite the possibility that there could be interactions between the regressors. For instance, alphabetically sorted authors could have different implications for each research discipline. Although it is sometimes common to sort authors alphabetically (66.1% of all the included articles with more than one author in statistics), authors were less often sorted alphabetically in Psychology (24.7%) or Economics (77.1%).

Including interaction terms for the above-mentioned effects, the interpretation of the results does not change. We report the estimated coefficients and ratios for this more complex model in Table 4. All included interaction terms were found to have a significant influence. Moreover, the AIC is smaller compared to the model reported in Table 3.

To control for the fact that the probability for name-sharing authors is increased with an increasing number of authors, we estimated a further model with only partial data. In particular, we only included articles that had exactly two authors. For this model (B), parameter estimates and ratios were shown in an analogous manner in Table 5. The results are in line with the results of the model described in Section 3.3, with a negative impact of authors sharing the same surnames, as well as more uncited articles of authors sharing the same surnames.

Table 1:

List of the included journals.

Field	Journals	Number of journals	Total number of articles	Average SJR	Average H	Average citations per document
Economics	Academy of Management Journal, Academy of Management Review, Accounting Review, Administrative Science Quarterly, American Economic Journal: Applied Economics, American Economic Journal: Economic Policy, American Economic Journal: Macroeconomics, American Economic Journal: Microeconomics, American Economic Review, Annual Review of Financial Economics, Econometrica, Experimental Economics, Journal of Accounting Research, Journal of Accounting and Economics, Journal of Economic Literature, Journal of Finance, Journal of Financial and Quantitative Analysis, Journal of Financial Economics, Journal of International Economics, Journal of Management, Journal of Monetary Economics, Journal of Political Economy, Journal of Supply Chain Management, Journal of the European Economic Association, Management Science, Quantitative Marketing and Economics, Quarterly Journal of Economics, Review of Economic Studies, Review of Financial Studies, Structural Equation Modeling, Theoretical Economics	31	44192	8.825	115.258	227.89
Psychology	Accounting, Organizations and Society, Annual Review of Psychology, Behaviour Research and Therapy, Biological Psychology, Child Development, Child Development Perspectives, Clinical Psychological Science, Clinical Psychology Review, Cognition, Cognitive Psychology, Current Directions in Psychological Science, Depression and Anxiety, Developmental Review, Developmental Science, Educational Psychologist, European Journal of Personality, European Review of Social Psychology, Evolution and Human Behavior, Frontiers in Behavioral Neuroscience, Frontiers in Human Neuroscience, Health Psychology Review, Journal of Abnormal Psychology, Journal of Applied Psychology, Journal of Child Psychology and Psychiatry and Allied Disciplines, Journal of Clinical Child and Adolescent Psychology, Journal of Consulting and Clinical Psychology, Journal of Consumer Psychology, Journal of Educational Measurement, Journal of Experimental Psychology: General, Journal of Memory and Language, Journal of Organizational Behavior, Journal of Personality and Social Psychology, Journal of Research in Crime and Delinquency, Journal of the American Academy of Child and Adolescent Psychiatry, Journal of Youth and Adolescence, Learning and Instruction, Learning and Memory, Memory and Cognition, Neuropsychology Review, Neuroscience and Biobehavioral Reviews, Organizational Behavior and Human Decision Processes, Personality and Social Psychology Bulletin, Personality and Social Psychology Review, Personnel Psychology, Perspectives on Psychological Science, Political Psychology, Psychological Bulletin, Psychological Medicine, Psychological Methods, Psychological Review, Psychological Science, Psychological Science in the Public Interest, Supplement, Psychotherapy and Psychosomatics, Research on Language and Social Interaction, Social Cognitive and Affective Neuroscience, Social Issues and Policy Review, Trends in Cognitive Sciences	57	106406	3.515	116.860	111.1802
Statistics	Annales de l'Institut Henri Poincare (B) Probability and Statistics, Annals of Applied Probability, Annals of Applied Statistics, Annals of Mathematics, Annals of Probability, Annals of Statistics, Annual Review of Statistics and Its Application, Biometrika, Biostatistics, Electronic Journal of Probability, Finance and Stochastics, Journal of Business and Economic Statistics, Journal of Computational and Graphical Statistics, Journal of Multivariate Analysis, Journal of Statistical Planning and Inference, Journal of Statistical Software, Journal of the American Statistical Association, Journal of the Royal Statistical Society: Series A: Statistics in Society, Journal of the Royal Statistical Society: Series B: Statistical Methodology, Journal of the Royal Statistical Society: Series C: Applied Statistics, Probability Theory and Related Fields, Scandinavian Journal of Statistics, Scientific data, Statistica Sinica, Statistical Science, Statistics and Computing, Test	27	45767	2.848	63.963	50.28357

Table 2: Descriptive statistics of selected covariates.

Variable	Freq. of 0	Min.	L.Q.	Median	Mean	U.Q.	Max.	St. Dev.
Citations	0.124	0	5	21	123.25	108	56424	482.86
Citations (> 0)	-	1	9	29	140.76	126	56424	513.63
Years since publ.	-	1	5	10	11.71	18	27	7.68
Title length	-	9	57	77	81.20	100	567	33.47
Number of authors	-	1	1	2	2.74	3	50	2.26
Single author	0.742	-	-	-	-	-	-	-
Author name-sharing	0.998	-	-	-	-	-	-	-

Table 3: Estimated coefficients $\hat{\beta}_i^z$ and $\hat{\beta}_i^c$ as well as odds ratios $\hat{r}_i^z$ and incident risk ratios $\hat{r}_i^c$ of a zero-inflated negative binomial regression model for citation counts. The zero-inflated effect as well as the count effect are significant for all introduced regressors and p-values are given in parentheses.

Variable	i	Zero-inflation coefficients $\hat{\beta}_i^z$	$\hat{r}_i^z$	Count coefficients $\hat{\beta}_i^c$	$\hat{r}_i^c$
Regressors					
Intercept	0	2.760		3.589	
		(< 0.0001)		(< 0.0001)	
Field of research: Psychology	1	0.368	1.445	-0.256	0.774
		(< 0.0001)		(< 0.0001)	
Field of research: Statistics	2	-0.662	0.516	-1.095	0.334
		(< 0.0001)		(< 0.0001)	
Years since publication: in full years	3	-0.052	0.949	0.072	1.074
		(< 0.0001)		(< 0.0001)	
Title length: number of characters in title	4	0.015	1.015	-0.001	0.999
		(< 0.0001)		(< 0.0001)	
Number of authors	5	-4.638	0.010	0.027	1.027
		(< 0.0001)		(< 0.0001)	
Alphabetically ordered authors: true	6	0.539	1.714	0.201	1.222
		(0.214)		(< 0.0001)	
Author name-sharing: existent	7	1.450	4.264	-0.220	0.803
		(0.031)		(0.001)	
Common author name: first author within top 100 surnames	8	-0.227	0.797	0.043	1.044
		(< 0.0001)		(< 0.0001)	
$\log(\hat{\theta})$				-0.835	
				(< 0.0001)	
Summary Statistics					
AIC				1771146	
$\exp(\log(\hat{\theta}))$				0.547	
LR (null model)				22313.31	

Table 4:

Estimated parameters $\hat{\beta}_{A,i}$ with odds ratios $\hat{r}^z_{A,i}$ or incidence risk ratios $\hat{r}^c_{A,i}$ of the zero-inflated negative binomial model for the first alternative model (A) with p-values in parentheses.

Variable	i	Zero-inflation coefficients $\hat{\beta}^z_{A,i}$	$\hat{r}^z_{A,i}$	Count coefficients $\hat{\beta}^c_{A,i}$	$\hat{r}^c_{A,i}$
Regressors					
Intercept	0	2.139		3.721	
		(< 0.001)		(< 0.001)	
Field of research: Psychology	1	0.364	1.439	-0.214	0.807
		(< 0.001)		(< 0.001)	
Field of research: Statistics	2	-0.731	0.481	-1.092	0.336
		(< 0.001)		(< 0.001)	
Years since publication: in full years	3	-0.052	0.950	0.067	1.076
		(< 0.001)		(< 0.001)	
Title length: number of characters in title	4	0.016	1.016	-0.001	0.999
		(< 0.001)		(< 0.001)	
Number of authors	5	-4.085	0.017	0.011	1.011
		(< 0.001)		(0.161)	
Single author (additional effect)	6	-		-0.288	0.750
				(< 0.001)	
Alphabetically ordered authors: true	7	-0.018	0.982	0.157	1.170
		(0.955)		(< 0.001)	
Author name-sharing: existent	8	1.476	4.377	-0.211	0.810
		(0.029)		(0.001)	
Common author name:	9	-0.238	0.788	0.042	1.042
first author within top 100 surnames		(< 0.001)		(0.002)	
Interaction: number of authors in Psychology	10	-		-0.014	0.986
		-		(0.068)	
Interaction: number of authors in Statistics	11	-		0.010	1.010
		-		(0.229)	
Interaction: alph. ordered authors in Psychology	12	-		-0.139	0.870
		-		(< 0.001)	
Interaction: alph. ordered authors in Statistics	13	-		-0.071	0.932
		-		(0.001)	
$\log(\hat{\theta})$				-0.604	
				(< 0.0001)	
Summary Statistics					
AIC				1770226	
$\exp(\log(\hat{\theta}))$				0.547	
LR (null model)				22778.4	

Table 5:
Estimated parameters $\hat{\beta}_{B,i}$ with odds ratios $\hat{r}^z_{B,i}$ and incidence risk ratios $\hat{r}^c_{B,i}$ of the zero-inflated negative binomial model for all articles of only two authors (alternative model B) and with p-values in parentheses.

Variable	i	Zero-inflation coefficients $\hat{\beta}^z_{B,i}$	$\hat{r}^z_{B,i}$	Count coefficients $\hat{\beta}^c_{B,i}$	$\hat{r}^c_{B,i}$
Regressors					
Intercept	0	-4.743		3.708	
		(< 0.001)		(< 0.001)	
Field of research: Psychology	1	-2.410	0.090	-0.126	0.881
		(< 0.001)		(< 0.001)	
Field of research: Statistics	2	-2.937	0.053	-0.997	0.369
		(< 0.001)		(< 0.001)	
Years since publication: in full years	3	-0.057	0.944	0.064	1.066
		(0.003)		(< 0.001)	
Title length: number of characters in title	4	0.030	1.031	-0.001	0.999
		(< 0.001)		(< 0.001)	
Number of authors	5	-	-	-	-
Alphabetically ordered authors: true	6	-1.044	0.352	0.205	1.228
		(< 0.001)		(< 0.001)	
Author name-sharing: existent	7	1.080	2.945	-0.176	0.839
		(0.072)		(0.006)	
Common author name: first author within top 100 surnames	8	-1.460	0.232	0.038	1.039
		(0.051)		(0.086)	
Interaction: number of authors in Psychology	9	-	-	-	-
Interaction: number of authors in Statistics	10	-	-	-	-
Interaction: alph. ordered authors in Psychology	11	-	-	-0.211	0.810
				(< 0.001)	
Interaction: alph. ordered authors in Statistics	12	-	-	-0.121	0.886
				(0.004)	
$\log(\hat{\theta})$				-0.561	
				(< 0.0001)	
Summary Statistics					
AIC				591036.7	
$\exp(\log(\hat{\theta}))$				0.571	
LR (null model)				5106.93	