This is a preprint of a paper accepted for publication in
Research Evaluation

From black box to white box at open access journals: Predictive validity of manuscript reviewing and editorial decisions at *Atmospheric Chemistry and Physics*

Lutz Bornmann,[*] Werner Marx,[†] Hermann Schier,[†] Andreas Thor,[§] and Hans-Dieter Daniel[*$]

[*] ETH Zurich, Professorship for Social Psychology and Research on Higher Education, Zähringerstr. 24, CH-8092 Zurich

[†] Max Planck Institute for Solid State Research, Heisenbergstraße 1, D-70569 Stuttgart

[§] University of Leipzig, Department of Computer Science, PF 100920, D-04009 Leipzig

[$] University of Zurich, Evaluation Office, Mühlegasse 21, CH-8001 Zurich

Address for correspondence:
Lutz Bornmann
ETH Zurich
Professorship for Social Psychology and Research on Higher Education
Zähringerstr. 24
CH-8092 Zurich
E-mail: bornmann@gess.ethz.ch

Abstract

More than 4500 open access (OA) journals have now become established in science. But doubts exist about the quality of the manuscript selection process for publication in these journals. In this study we investigate the quality of the selection process of an OA journal, taking as an example the journal *Atmospheric Chemistry and Physics* (ACP). ACP is working with a new system of public peer review. We examine the predictive validity of the ACP peer review system – namely, whether the process selects the best of the manuscripts submitted. We have data for 1111 manuscripts that went through the complete ACP selection process in the years 2001 to 2006. The predictive validity was investigated on the basis of citation counts for the later published manuscripts. The results of the citation analysis confirm the predictive validity of the reviewers' ratings and the editorial decisions at ACP: Both covary with citation counts for the published manuscripts.

Key words: Public peer review, Open access, Predictive validity

1. Introduction

"Since so much hinges on peer review and it is so central to what and where things are published, it is essential that it is carried out well and professionally" (Hames, 2007, p. 2). The most serious limitation of research on the predictive validity of reviewers' recommendations and editorial decisions is the very small number of studies. Up to now only a few studies have conducted analyses that examine citation counts from individual papers as the basis for assessing predictive validity in peer review. Research in this area is extremely labor-intensive, since a validity test requires information and citation counts regarding the fate of rejected manuscripts (Bornstein, 1991). The editor of the *Journal of Clinical Investigation* (Wilson, 1978) undertook his own investigation into the question of predictive validity. Daniel (1993) and Bornmann and Daniel (2008a, 2008c) investigated the peer review process of *Angewandte Chemie International Edition* (AC-IE), and Opthof *et al* (2000) did the same for *Cardiovascular Research*. McDonald *et al* (2009) examined the predictive validity of the editorial decisions for the *American Journal of Neuroradiology*. All five studies confirmed that the editorial decisions (acceptance or rejection) for the various journals appear to reflect a rather high degree of predictive validity, if citation counts are employed as validity criteria.

More than 4500 (interactive) open access (OA) journals have now become established in science that work either with the traditional peer review system or with the 'new' system of public peer review (http://www.doaj.org/). The strongest reservation about OA journals is doubt as to whether they achieve sufficient quality control (Joint Information Systems Committee, 2004). "In the open-access business model, it is widely accepted that authors (or their funding agencies or universities) pay. This means that … the earnings of the journal are directly dependent on the number of articles published. Only fools believe that editors wouldn't then tend towards acceptance of a manuscript in the many borderline cases" (Gölitz, 2010, p. 4). According to Taylor *et al* (2008) "one may argue that editors of OA … journals, pressured by their commercial employers, may be forced to accept articles of lower quality in order to increase the number published and therefore the journal's income." It was, for example, recently discovered that Merck – a global pharmaceutical and chemical company – "paid an undisclosed sum to Elsevier [Amsterdam, The Netherlands] to produce several volumes of a publication that had the look of a peer-reviewed medical journal, but contained only reprinted or summarized articles – most of which presented data favorable to Merck products – that appeared to act solely as marketing tools with no disclosure of company sponsorship" (Grant, 2009, para. 1).

In this study the quality of the manuscript selection process of an OA journal is investigated for the first time (to our knowledge), taking the journal *Atmospheric Chemistry and Physics* (ACP) as an example. We examined whether the ACP peer review system in fact selects the 'best' manuscripts for publication from all manuscripts submitted. To do this, in a first step of the analysis we compared the citation impact of papers that received a positive rating by reviewers and were published in ACP and papers that received a negative review and were rejected for publication in ACP but were later published elsewhere. As at ACP comparatively few manuscripts are rejected for publication, in a second step of the analysis we examined the correlation between reviewers' ratings (on which the editorial decisions to accept or reject are based) and citation counts. As the number of citations of a publication reflects the international impact of the reported research, and for lack of other operationalizable indicators, it is a common approach in peer review research to evaluate the success of a peer review process on the basis of the citation count of the reviewed manuscripts (Bornmann, in press). According to Jennings (2006) "the most important question is how accurately the peer review system predicts the longer-term judgments of the scientific community."

Against the backdrop of these and similar statements, scientific judgments on manuscripts are said to show predictive validity, if favorable reviewers' ratings or acceptance decisions, respectively, are statistically significantly associated with increasing citation counts.

2. Methods

Manuscript review at ACP

ACP was launched in September 2001. It is produced and published by the European Geosciences Union (EGU) (www.copernicus.org/EGU/EGU.html) and the Copernicus Society (www.copernicus.org). ACP is freely accessible via the Internet (www.atmos-chem-phys.org). ACP has a two-stage publication process, with a 'new' peer review process consisting in public peer review and interactive discussion (Koop and Pöschl, 2006; Pöschl, 2004) that is described at the ACP website as follows: In the first stage, manuscripts that pass a rapid pre-screening process (access review) are immediately published as 'discussion papers' on the journal's *Atmospheric Chemistry and Physics Discussions* (ACPD) website. These discussion papers are then made available for 'interactive public discussion,' during which the comments of designated reviewers (usually, the reviewers that already conducted the access review), additional comments by other interested members of the scientific community, and the authors' replies are published alongside the discussion paper.

During the discussion phase, the designated reviewers are asked to answer four questions concerning the ACP principal manuscript evaluation criteria (see http://www.atmospheric-chemistry-and-physics.net/review/ms_evaluation_criteria.html). The reviewers are asked to rate the manuscripts on scientific quality, scientific significance, presentation quality, and worthiness of publication. The scientific quality criterion asks, for example: "Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?" The response categories for the questions are: (1) excellent, (2) good, (3) fair, and (4) poor. After the end of the discussion phase every author has the opportunity to submit a revised manuscript taking into account the reviewers' comments and the comments of interested members of the scientific community. Based on the revised manuscript and in view of the access peer review and interactive public discussion, the editor accepts or rejects the revised manuscript for publication in ACP. For this decision, further external reviewers may be asked to review the revision, if needed.

Database for the present study

For the investigation of peer review at ACP we had data for 1111 manuscripts that went through the complete ACP selection process in the years 2001 to 2006. These manuscripts reached one of the following final statuses: 958 (86%) were published in ACPD and ACP, 74 (7%) were published in ACPD but not in ACP (here, the editor rejected the revised manuscript), and 79 (7%) were not published in either ACPD or ACP (these manuscripts were rejected during the access review). Some of the manuscripts submitted to ACP but not published there (because they were rejected during the access review, for example) were submitted by the authors, as described in the following, to another journal and published there. According to Schultz (2010), there are two reasons for the high publication rate of submissions to ACP (see also Pöschl, 2010): By using the public peer review and interactive discussion, (1) ACP can expect a high average quality of submitted manuscripts, and (2) ACP works harder than journals working with the traditional peer review to keep and improve the submissions.

Reviewers' ratings on the scientific quality of the manuscripts were available for 552 (55%) of the 1008 manuscripts that were subject to full peer review and interactive public discussion. This reduction in number is due to the fact that the publisher has stored the ratings electronically only since 2004. In the evaluation of predictive validity in this study we included the ratings only for those manuscripts that were later published in ACP (*n*=496). In this way, factors were kept constant that could have an undesired influence on the investigation of the association between reviewers' ratings and citation counts (for example, the visibility of the journal publishing the manuscript). Of these 496 manuscripts, 17% (*n*=84) had one review, 63% (*n*=313) had two, 17% (*n*=83) had three, 3% (*n*=14) had four, and 2 manuscripts had five independent reviews. For the statistical analysis, for each manuscript the median of the independent ratings for the scientific quality was computed. According to Thorngate *et al* (2009), the average error in ratings decreases with an increasing number of raters.

Conducting of citation analysis

Citation counts are attractive raw data for the evaluation of research output: They are "unobtrusive measures that do not require the cooperation of a respondent and do not themselves contaminate the response (i.e., they are non-reactive)" (Smith, 1981, p. 84). Although citations have been controversial as a measure of quality and scientific progress (see an overview of the discussion in Bornmann and Daniel, 2008d), they are accepted as a measure of scientific impact and thus as a partial aspect of scientific quality (Martin and Irvine, 1983). For Lindsey (1989), citations are "our most *reliable* convenient measure of quality in science – a measure that will continue to be widely used" (p. 201). The findings of Bornmann and Daniel (2008b) suggested "that the more an article is cited, the more intensively its content is used by the citing scientists" (p. 37).

For manuscripts published in ACP, ACPD, or elsewhere, we determined the total number of citations up to the end of 2008 and the number of citations for a fixed time window of three years including the publication year. "Fixed citation windows are a standard method in bibliometric analysis, in order to give equal time spans for citation to articles published in different years, or at different times in the same year" (Craig *et al*, 2007, p. 243). According to Harnad (2007), the reliability of the results of citation analysis can be tested by using two different citation windows (here: a fixed three-year window, and from publication up to the end of 2008). In the citation search we included self-citations, because (1) it is not expected that the number of self-citations varies systematically for the manuscripts published in ACP, ACPD, or elsewhere, and (2) the number of self-citations of a publication can be modeled in the multiple regression analysis (the results of which are reported in the following) using the number of authors of a manuscript. As Herbertz (1995) shows, a greater number of authors is associated with a greater number of self-citations of a publication (see also Leimu and Koricheva, 2005).

The citation analyses for the present study were conducted based on the Science Citation Index (SCI) (Thomson Reuters, Philadelphia, PA, USA), Chemical Abstracts (CA) (Chemical Abstracts Services, Columbus, Ohio, USA), Scopus (Elsevier, Amsterdam, The Netherlands), and Google Scholar (GS) (Google, Inc., headquartered in Mountain View, California). The SCI includes multidisciplinary data from journals in the sciences (see http://isiwebofknowledge.com). These journals also form the database for Scopus and GS, but both include a greater number of journals. GS does not search only peer-reviewed research journals, but searches lots of non-traditional sources; including institutional repositories preprint archives and conference proceedings (Bornmann *et al*, 2009). CA is a comprehensive database of publicly disclosed research in chemistry and related sciences (see http://www.cas.org/). A study by Whitley (2002) comparing citation searching in Web of Science (WoS) (Thomson Reuters; WoS provides access to the SCI) and CA showed that the two indices lead to different results. According to Whitley (2002), CA misses an estimated 17% of the citations found only in WoS. Conversely, researchers using only WoS to search citations miss an estimated 23% covered only in CA. Vieira and Gomes (2009) examined the coverage achieved by WoS and by Scopus: “The general conclusion is that about 2/3 of the documents referenced in any of the two databases may be found in both databases while a fringe of 1/3 are only referenced in one or the other” (p. 587).

Manuscripts published in ACPD are included as documents in the source index of

Scopus (however, not all volumes that have been published up to now, see here also Jacso, 2009a) but not in CA or SCI. However, their citations are searchable via a method that tallies with the Cited Reference Search in WoS. For a manuscript, the frequency of the different variants of the journal title of ACPD (for example, Atm Chem Phys Disc, Atm Chem Phys Discus, Atmos Chem Phys Disc) in combination with the publication year are searched within all of the references (citations) contained in the database and restricted to the specific time window. If a manuscript in our dataset was published not only in ACPD but also in ACP (or another journal), the citation counts for both publications are added up. The addition of both citation counts was done in view of the fact that double count citations (that is, citation of both publications of a manuscript within one paper) are very rare (see here also Bloom, 2006). (For about 10 ACP papers (of about 1000) in our dataset this resulted in inclusion of citations of the ACPD manuscript but of no citations of the ACP manuscript in the added up citation counts for the three-year window. For these papers a three-year citation window was available for the ACPD manuscript but not for the ACP manuscript.)

Checking for double count citations was carried out using a recently developed routine for macro programming of the Messenger command language from STN International (Eggenstein-Leopoldshafen, Germany). This allowed examination of the number of double count citations of the 958 individual ACP papers with the corresponding papers in ACPD in SCI up to the present. Only 18 true double count citations were found where an ACP paper was cited together with the corresponding paper published in ACPD. In addition, we did a manual check of the number of double count citations for the complete ACP publication year 2004 as an example: For 2004 SCI shows 174 ACP papers as source items. The intersection of the 2320 papers citing these ACP papers with the 371 papers citing the corresponding ACPD papers was 90 citing papers. In these 90 citing papers, at least one ACP paper from 2004 was cited together with an ACPD paper from 2004. A manual check of the citations of the ACP and ACPD papers in the citing papers revealed only 3 true double count citations. As using the two methods the citations across the complete time period were included, the number of double count citations for a three-year window is smaller. Usually, the authors cited different papers and not corresponding ACP and ACPD papers.

Factors with a general influence on citation counts

Bibliometric studies have demonstrated that several factors have a *general* influence on citation counts. By considering these factors in the statistical analysis, it becomes possible to establish the adjusted covariation between editorial decisions or reviewers' ratings, respectively, and citation counts of manuscripts published in ACP, ACPD, or elsewhere.

The probability of citations may be influenced by the number of authors (Vieira and Gomes, 2010), the number of pages in a journal (Bornmann and Daniel, 2007), and the field or discipline to which the manuscript can be assigned (Bornmann and Daniel, 2008d). The Journal Impact Factor (JIF) was not considered as a predictive factor in the statistical analysis because "article citation rates determine the journal impact factor, not vice versa" (Seglen, 1997). JIFs are published by Thomson Reuters in the Journal Citation Reports (JCR) and are a measure of the 'average' response of the scientific community to a paper in a journal (Bornmann *et al*, 2007). For Leimu and Koricheva (2005) it is a widespread belief "that publication in a high-impact journal might by itself enhance the citation rate of an article by increasing its visibility or persuasiveness of the arguments presented" (p. 29). Their study results "do not support this 'journal effect' hypothesis, because there was considerable variation in citation rates, especially for papers published in high-impact journals" (Leimu and Koricheva, 2005, p. 29).

We performed multiple regression analyses, which reveal the factors that exert a primary influence on a certain outcome. The coefficients in the regression model, called 'partial' regression coefficients (Rabe-Hesketh and Everitt, 2004), represent the effects of each factor, controlling for all other factors in the model. Since the skewness of citation counts suggests the use of a negative binomial specification (Allison *et al*, 1982), we calculated a Negative Binomial Regression Model (NBRM, Long and Freese, 2006, section 8.3) (see also Bornmann *et al*, 2008). The citation counts for the manuscripts published in ACP, ACPD, or elsewhere enter into the NBRM as a dependent variable. As we performed the citation search for the manuscripts in SCI, CA, Scopus, and GS for two different citation windows (except GS), we calculated several different models. We performed bootstrapping to improve the estimations of the standard errors of the model (Cameron and Trivedi, 1998), since the group of manuscripts that were not published in ACP but published elsewhere is comparatively small compared to the manuscripts published in ACP.

3. Results

The fate of manuscripts that were not published in ACP

The search for the fate of the manuscripts that were not published in ACP (*n*=153) was conducted using two research literature databases, WoS and CA. Two Ph.D. environmental research scientists carried out the search. The results of the investigation revealed that of the 153 manuscripts, 38 (25%) were published in other journals. No publication information was found for 115 (75%) manuscripts, whereby 70 of the 115 manuscripts (61%) were published in ACPD. Other studies on the fate of manuscripts that were rejected by a journal reported percentages ranging from 28% to nearly 85% for manuscripts later published elsewhere (Weller, 2002), whereby the journals examined do not work with a two-stage publication process as does ACP. For manuscripts rejected by AC-IE at the beginning of the year 2000, Bornmann and Daniel (2008a) determined a percentage of 95%.

Where investigation of the fate of manuscripts that were not published in ACP established that a manuscript had been subsequently published elsewhere, the two Ph.D. environmental scientists determined the extent of changes that had been made to the manuscript. Table 1 shows the results of the search for the manuscripts that were not published in ACP. The 38 manuscripts that were published as contributions in other journals were published in 25 different journals (see Table 1) within a time period of five years (that is, between 2005 and 2009). Six manuscripts were published in the *Journal of Geophysical Research*; three manuscripts were published in *Geophysical Research Letters*. The other 23 journals each published one or two manuscripts. As Table 1 further shows (see Total), on average (median) no changes or minor changes were made to the manuscripts. A similar result was reported by two other studies: Relman (1978) reported for *The New England Journal of Medicine*: "In approximately 80 percent of … cases, the manuscript had not been changed appreciably. Of the remaining 20 percent, most were altered only moderately" (p. 58). For AC-IE, Bornmann and Daniel (2008a) found: "Of 1021 Communications rejected for publication by *Angewandte Chemie*, 959 were published as contributions (93.9%) in other journals, seven as patents, and two as contributions to anthologies … No alterations or only minor alterations were made to approximately three-quarters of the rejected Communications for publication elsewhere" (p. 7175).

Journal Impact Factors of the journals in which manuscripts were later published

In the following, we examine the predictive validity of the peer review process at ACP, based on the mean citation counts of manuscripts published in ACP, ACPD, or elsewhere and on the JIFs of the journals in which manuscripts were later published. Published annually by Thomson Reuters, the 3-year Impact Factor (JIF_3) is the quotient of citations and number of citable items: It is determined based on the papers published in a journal in a two-year period and their citations received in the year thereafter. The number of times that the papers were cited is then divided by the number of citable items published in the previous two years (Garfield, 2006). The JIF_3 refers to a period of three years, and the 5-year Impact Factor (JIF_5) is based on a period of five years (Jacso, 2009b).

According to the Thomson Reuters' JCR for the year 2008 (see Table 1), the JIFs of the journals that published the rejected manuscripts ranged from less than 1 (for example, *Acta Chemica Slovenica* and *Advances in Atmospheric Sciences*) to 4.215 (JIF_3) or 4.491 (JIF_5) (*Crystal Growth & Design*), respectively. As measured by these JIFs, the ACP editorial decisions to accept or rejects manuscripts are highly valid. All of the 38 manuscripts in Table 1 were published in a journal having a lower JIF 2008 than ACP (JIF_3 = 4.927, JIF_5 = 4.904). A similar result was found by Daniel (1993) and Bornmann and Daniel (2008a) for manuscripts rejected by AC-IE in the year 1984 and at the beginning of 2000. These findings confirm Cronin and McKenzie's (1992) general observation that manuscripts that are rejected by prestigious journals having high JIFs are usually later submitted to (and published by) journals having lower JIFs: "It is widely recognised that there is an informal journal pecking order in almost every discipline, and that a manuscript rejected by a high-ranking journal will often be re-submitted to one of lesser repute" (p. 310).

Comparison of accepted or rejected but published elsewhere manuscripts

Median citation counts

The JIF is only a very rough measure for determining predictive validity, because all of the contributions in a journal are characterized by an average value (Braun et al, 2007). It thus overestimates the citations of the top cited papers while de-emphasizing the number of citations of papers that are not or infrequently cited. For this reason, going beyond JIFs, we determined how frequently the manuscripts published in ACP, ACPD, or elsewhere were cited Table 2 shows the mean number of citations found in CA, SCI, Scopus, and GS for two different citation windows. (With GS, it is not possible to search the number of citations for a fixed time window because the publication year of the citing publications is not available in every case.) The medians are printed in bold in the table, as the median is not affected by outliers, unlike the arithmetic mean. The results are shown for three manuscript groups. Group 1 manuscripts are manuscripts that were published in ACP (and ACPD) (accepted manuscripts). Group 2 manuscripts were rejected after the discussion phase and published either only in ACPD or in ACPD and elsewhere. Group 3 manuscripts were rejected in the access review and published elsewhere.

As the results in Table 2 show, independently of the literature database in which the citation search was conducted and the citation time window, group 1 manuscripts were cited more frequently on average than those in group 3. For example, manuscripts that were published in ACP and ACPD (group 1) were cited, according to the SCI, on average 10 times (median) up to the end of 2008; manuscripts that were published in neither ACP nor ACPD but published elsewhere (group 3) were cited on average 1.5 times (median) up to the end of 2008. The results also show that the manuscripts in group 1 were cited clearly more frequently than those that were published in ACPD only or in ACPD and elsewhere (group 2). In contrast, hardly any differences were found between the median citation counts of group 2 and group 3 manuscripts.

Negative binomial regression models

While the medians in Table 2 suggest that ACP indeed publishes ‘the better’ manuscripts among the submissions, factors other than their scientific contribution to the development of an important area of research could in principle have been responsible for the higher citation counts. We performed seven NBRMs (with citations found in CA, SCI, Scopus, and GS for two different citation windows as outcome variables), which reveal these factors. As only those cases could be included in the statistical analyses that had no missing values for the dependent and independent variables that were entered into the NBRMs, the models had to be calculated with a reduced sample size. As a result, between 955 (model G) and 989 (models A and C) of the manuscripts could be included. As the reduction of the sample sizes resulted in an insufficient number of cases in group 2 (manuscripts published in ACPD only or in ACPD and elsewhere) and group 3 (manuscripts published in neither ACP nor ACPD but published elsewhere), the two groups were combined into one group. This group contained all manuscripts that were not published in ACP but published elsewhere.

Table 3 shows a description of the independent variables that were included in the NBRMs. In addition to the editorial decision (to publish or not publish in ACP), the models take into account the number of co-authors, the number of pages in the manuscripts, and the research subfield. CAS categorizes publications in different chemical 'sections,' each section covering only one broad area of scientific inquiry. Normally, each abstract in CA appears in only one CA section, based on the most important aspect of the publication. As Table 3 shows, about one-half of the manuscripts that were included in the NBRMs were assigned to the section 'Air Pollution and Industrial Hygiene' and a further 21% to the section 'Mineralogical and Geological Chemistry.' As the manuscripts that were assigned by CAS to other sections (2% of the manuscripts) usually did not appear in the dataset more frequently than twice, for the statistical analysis they were grouped together in a category called 'Other Section.' For 30% of the manuscripts no section in CA could be found. As WoS and Scopus do not assess publications by content (Neuhaus and Daniel, 2008), these databases could not be used as an alternative to CA.

In models A, C, E, and G the publication year of each manuscript was included in the models as exposure time (Long and Freese, 2006, pp. 370-372). By using the exposure option provided in the statistical package Stata (StataCorp., 2008), the amount of time that an article is 'at risk' of being cited is considered. In models B, D, and F this option is not needed, as the individual citation counts refer to the same citation window of three years.

In Table 4 the results of models A to G for predicting citation counts for manuscripts published in ACP or elsewhere are similar. As for the variable 'editorial decision,' the models yield a statistically significant greater number of citations for manuscripts published in ACP than manuscripts published elsewhere. In the light of CA, SCI, Scopus, and GS citations for two citation windows, therefore, ACP editors were able to accomplish the difficult task of publishing the 'best' manuscripts among submissions. Furthermore, in all of the models, statistically significant effects could be found in the expected directions for two factors that in bibliometric studies have been demonstrated to have a *general* influence on citation counts: A higher number of co-authors and a higher number of pages are associated with a greater number of citations. The results for the influence of the CA section are inconsistent: In four models (A, B, C, and E) a statistically significant correlation could be found, but in three models (D, F and G) the correlation was not significant statistically.

Comparison of manuscripts with different reviewers' ratings

Median citation counts

More than 90% of the manuscripts that were included in the above citation analysis for accepted or rejected but published elsewhere manuscripts belong to only one group (group 1, published in ACP and ACPD). For this reason, for further examination of the predictive validity of the ACP review process we undertook a differentiation within this group on the basis of the reviewers' ratings (median reviewers' ratings on the scientific quality of the manuscripts). Since according to Thorngate *et al* (2009) assessment of the importance or significance of a scientific work is more a matter of taste on the part of the reviewer, the analysis of the predictive validity of the reviewers' ratings is not based on the question asked of the reviewers concerning the importance but rather on the question concerning the quality of the manuscript (see the Methods section). In this analysis, we examined to what extent the ratings for group 1 manuscripts received during the discussion phase of the peer review process correlated with the citation counts for the manuscripts later published in ACP. Table 5 shows the citation counts for the manuscripts sorted by different median ratings. As the median citation counts for the manuscripts with different ratings show, with some exceptions, better median ratings are correlated with an increase in the median citation count.

Negative binomial regression models

Since also citation counts for ACP papers in Table 5 can be affected by factors other than the papers' scientific contribution to the development of an important area of research, a further six NBRMs including the reviewers' median ratings as independent variables were computed. Table 6 shows a description of the dependent and independent variables that were included in the model. As the same 496 manuscripts were included in models A to F (each with different citation counts from the three literature databases), Table 6 shows different values for the dependent variable, the citations, but not for the independent variables (such as the reviewers' median ratings). In model G a somewhat reduced sample size could be included (n=482).

Table 7 shows the results of the NBRMs. All models show a statistically significant correlation between the reviewers' median ratings and the citation counts: The higher the average reviewers' rating on quality of a manuscript, the higher the number of citations that the paper published later in ACP had. This result is in agreement with the results of the NBRMs in Table 4 on the editorial decision. As for the other independent variables that were included in the models (factors that have a general influence on citation counts), the results in Table 7 differ somewhat from the results in Table 4. Only seven parameter estimates in Table 7 are statistically significant. When interpreting this difference, however, it should be taken into account that in the models in Table 7 fewer manuscripts could be included than in the models in Table 4, and it is thus more difficult to insure that group differences were not due to chance (Cohen, 1988).

4. Discussion

Many OA journals come into being in recent years. It is hoped that unrestricted access to scientific publications will have a positive effect on scientific progress: According to Borgman (2007), "scholarship is a cumulative process, and its success depends on wide and rapid dissemination of new knowledge so that findings can be discarded if they are unreliable or built on if they are confirmed. Society overall benefits from the open exchange of ideas within the scholarly community" (p. 35). Some of the OA journals are using public or open peer review, for one, in the interest of higher quality submissions: "Open review has the advantage of speeding and democratizing reviewing, and could result in better manuscripts being submitted" (Borgman, 2007, p. 61). Furthermore, "reviewers would be more tactful and constructive" (DeCoursey, 2006). And for another, "there is a widely held suspicion (certainly amongst commercial publishers and to a lesser extent amongst authors) that articles in … OA journals are less well peer-reviewed than their counterparts in toll-access journals. This perception has two roots; firstly, as … OA journals are new, they have not yet had a chance to attain high status, and secondly, there is a feeling that because income depends on the number of accepted articles, the editors will be under pressure to accept poor quality manuscripts to keep the income stream up" (Oppenheim, 2008, p. 582).

Contrary to those fears, the results of this study show – in agreement with the results on various closed peer review systems of traditional journals mentioned in the introduction to this paper – that in the journal examined here, public peer review is able to assess the quality of manuscripts 'validly' and to select the 'best' manuscripts among the manuscripts submitted. The results of the citation analysis confirm the predictive validity of the editorial decisions and reviewers' ratings: They correlate statistically significantly with citation counts. When interpreting these results, however, it should be taken into consideration that the ACP peer review system, through the high acceptance rate among submissions, in many cases exercises a different function than the peer review system at many traditional journals: It is more about improving manuscripts prior to publication than about selecting among submissions. In the words of Shashock (2005), journals like *Science* or *Nature* skim off the cream and discard everything else among the submissions. ACP, in contrast, in the first review step screens out unsuitable manuscripts only and eliminates them from the further selection process. Through the use of public peer review in the second review step, a large part of the manuscripts that in the access review were deemed potentially suitable for publication in ACP are published after varying degrees of revision.

Since the number of OA journals can be expected to increase in coming years, future studies on predictive validity should examine in particular their peer review systems. Here, studies are needed that investigate not only the selection function, as in this study, but also the improvement function of peer review.

Acknowledgements

We would like to thank Dr. Hanna Joos (at the Institute for Atmospheric and Climate Science of the ETH Zurich, Switzerland) and Dr. Hanna Herich (at EMPA – a research institution within the ETH Domain) for the investigation of the manuscripts rejected by *Atmospheric Chemistry and Physics* and published elsewhere. The research project, which is investigating quality assurance of interactive open access journals, is supported by a grant from the Max Planck Society (Munich, Germany). We thank Dr. Ulrich Pöschl, Chief Executive Editor of *Atmospheric Chemistry and Physics*, the Editorial Board of *Atmospheric Chemistry and Physics*, and Copernicus Publications (Katlenburg-Lindau, Germany) for permission to conduct the evaluation of the selection process of the journal, and thank the members of Copernicus Systems + Technology (Berlin, Germany) for their generous technical support during the carrying out of the study.

References

Allison, P D, J S Long and T K Krauze (1982), 'Cumulative advantage and inequality in science', *American Sociological Review*, 47(5), pages 615-625.

Bloom, T. (2006). *Systems: online frontiers of the peer-reviewed literature. The Internet is allowing much more interactive science publishing*. Retrieved July 6, 2006, from http://www.nature.com/nature/peerreview/debate/nature05030.html

Borgman, C L (2007), *Scholarship in the digital age. Information, infrastructure, and the Internet* (MIT Press, Cambridge, MA, USA).

Bornmann, L (in press), 'Scientific peer review', *Annual Review of Information Science and Technology*.

Bornmann, L and H-D Daniel (2007), 'Multiple publication on a single research study: does it pay? The influence of number of research articles on total citation counts in biomedicine', *Journal of the American Society for Information Science and Technology*, 58(8), pages 1100-1107.

Bornmann, L and H-D Daniel (2008a), 'The effectiveness of the peer review process: inter-referee agreement and predictive validity of manuscript refereeing at *Angewandte Chemie*', *Angewandte Chemie International Edition*, 47(38), pages 7173-7178.

Bornmann, L and H-D Daniel (2008b), 'Functional use of frequently and infrequently cited articles in citing publications. A content analysis of citations to articles with low and high citation counts', *European Science Editing*, 34(2), pages 35-38.

Bornmann, L and H-D Daniel (2008c), 'Selecting manuscripts for a high impact journal through peer review: a citation analysis of Communications that were accepted by *Angewandte Chemie International Edition*, or rejected but published elsewhere', *Journal of the American Society for Information Science and Technology*, 59(11), pages 1841-1852.

Bornmann, L and H-D Daniel (2008d), 'What do citation counts measure? A review of studies on citing behavior', *Journal of Documentation*, 64(1), pages 45-80.

Bornmann, L, L Leydesdorff and W Marx (2007), 'Citation environment of *Angewandte Chemie*', *CHIMIA*, 61(3), pages 104-109.

Bornmann, L, W Marx, H Schier, E Rahm, A Thor and H D Daniel (2009), 'Convergent validity of bibliometric Google Scholar data in the field of chemistry. Citation counts for papers that were accepted by *Angewandte Chemie International Edition* or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts', *Journal of Informetrics*, 3(1), pages 27-35.

Bornmann, L, R Mutz, C Neuhaus and H-D Daniel (2008), 'Use of citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results', *Ethics in Science and Environmental Politics*, 8, pages 93-102.

Bornstein, R F (1991), 'The predictive validity of peer-review: a neglected issue', *Behavioral and Brain Sciences*, 14(1), pages 138-139.

Braun, T, I Dióspatonyi, S Zsindely and E Zádor (2007), 'Gatekeeper index versus impact factor of science journals', *Scientometrics*, 71(3), pages 541-543.

Cameron, A C and P K Trivedi (1998), *Regression analysis of count data* (Cambridge University Press, Cambridge, UK).

Cohen, J (1988), *Statistical power analysis for the behavioral sciences* (Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, USA).

Craig, I D, A M Plume, M E McVeigh, J Pringle and M Amin (2007), 'Do open access articles have greater citation impact? A critical review of the literature', *Journal of Informetrics*, 1(3), pages 239-248.

Cronin, B and G McKenzie (1992), 'The trajectory of rejection', *Journal of Documentation*, 48(3), pages 310-317.
Daniel, H-D (1993), *Guardians of science. Fairness and reliability of peer review* (Wiley-VCH, Weinheim, Germany).
DeCoursey, T. (2006). *Perspective: The pros and cons of open peer review. Should authors be told who their reviewers are?* Retrieved June 20, 2006, from http://www.nature.com/nature/peerreview/debate/nature04991.html
Garfield, E (2006), 'The history and meaning of the Journal Impact Factor', *Journal of the American Medical Association*, 295(1), pages 90-93.
Gölitz, P (2010), 'Twitter, Facebook, and Open Access ...', *Angewandte Chemie International Edition*, 49(1), pages 4-6.
Grant, B. (2009). *Merck published fake journal*. Retrieved March 25, 2010, from http://www.the-scientist.com/blog/display/55671/
Hames, I (2007), *Peer review and manuscript management of scientific journals: guidelines for good practice* (Blackwell, Oxford, UK).
Harnad, S (2007), 'Open Access scientometrics and the UK Research Assessment Exercise', in D Torres-Salinas and H F Moed (Editors), *Proceedings of the 11th Conference of the International Society for Scientometrics and Informetrics* (Spanish Research Council (CSIC), Madrid, Spain), pages 27-33.
Herbertz, H (1995), 'Does it pay to cooperate? A bibliometric case-study in molecular-biology', *Scientometrics*, 33(1), pages 117-122.
Jacso, P (2009a), 'Database source coverage: hypes, vital signs and reality checks', *Online Information Review*, 33(5), pages 997-1007.
Jacso, P (2009b), 'Five-year impact factor data in the Journal Citation Reports', *Online Information Review*, 33(3), pages 603-614.
Jennings, C G. (2006). *Quality and value: the true purpose of peer review. What you can't measure, you can't manage: the need for quantitative indicators in peer review*. Retrieved July 6, 2006, from http://www.nature.com/nature/peerreview/debate/nature05032.html
Joint Information Systems Committee. (2004). *Journal authors survey report*. Truro, UK: Key Perspectives Ltd.
Koop, T and U Pöschl. (2006). *Systems: an open, two-stage peer-review journal. The editors of Atmospheric Chemistry and Physics explain their journal's approach.* Retrieved 26 June 2006, from http://www.nature.com/nature/peerreview/debate/nature04988.html
Leimu, R and J Koricheva (2005), 'What determines the citation frequency of ecological papers?', *Trends in Ecology & Evolution*, 20(1), pages 28-32.
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), pages 189-203.
Long, J S and J Freese (2006), *Regression models for categorical dependent variables using Stata* (Stata Press, Stata Corporation, College Station, TX, USA).
Marsh, H W, L Bornmann, R Mutz, H D Daniel and A O'Mara (2009), 'Gender effects in the peer reviews of grant proposals: a comprehensive meta-analysis comparing traditional and multilevel approaches', *Review of Educational Research*, 79(3), pages 1290-1326.
Martin, B R and J Irvine (1983), 'Assessing basic research - some partial indicators of scientific progress in radio astronomy', *Research Policy*, 12(2), pages 61-90.
McDonald, R J, H J Cloft and D F Kallmes (2009), 'Fate of manuscripts previously rejected by the *American Journal of Neuroradiology*: a follow-up analysis', *American Journal of Neuroradiology*, 30(2), pages 253-256.
Neuhaus, C and H-D Daniel (2008), 'Data sources for performing citation analysis - an overview', *Journal of Documentation*, 64(2), pages 193-210.
Oppenheim, C (2008), 'Electronic scholarly publishing and open access', *Journal of*

Information Science, 34(4), pages 577-590.
Opthof, T, F Furstner, M van Geer and R Coronel (2000), 'Regrets or no regrets? No regrets! The fate of rejected manuscripts', *Cardiovascular Research*, 45(1), pages 255-258.
Pöschl, U (2004), 'Interactive journal concept for improved scientific publishing and quality assurance', *Learned Publishing*, 17(2), pages 105-113.
Pöschl, U (2010), 'Interactive open access publishing and peer review: the effectiveness and perspectives of transparency and self-regulation in scientific communication and evaluation', *Liber Quarterly*, 19(3/4), pages 293-314.
Rabe-Hesketh, S and B Everitt (2004), *A handbook of statistical analyses using Stata* (Chapman & Hall/CRC, Boca Raton, FL, USA).
Relman, A S (1978), 'Are journals really quality filters?', in W Goffman, J T Bruer and K S Warren (Editors), *Coping with the biomedical literature explosion: a qualitative approach* (The Rockefeller Foundation, New York, NY, USA), pages 54-60.
Schultz, D M (2010), 'Rejection rates for journals publishing atmospheric science', *Bulletin of the American Meteorological Society*, 91(2), pages 231-243.
Seglen, P O (1997), 'Why the impact factor of journals should not be used for evaluating research', *British Medical Journal*, 314(7079), pages 498-502.
Shashok, K (2005), 'Standardization vs diversity: how can we push peer review research forward?', *Medscape General Medicine*, 7(1), page 11.
Smith, L C (1981), 'Citation analysis', *Library Trends*, 30(1), pages 83-106.
StataCorp. (2008), *Stata statistical software: release 11* (Stata Corporation, College Station, TX, USA).
Taylor, M, P Perakakis and V Trachana (2008), 'The siege of science', *Ethics in Science and Environmental Politics*, 8, pages 17-40.
Thorngate, W, R M Dawes and M Foddy (2009), *Judging merit* (Psychology Press, New York, NY, USA).
Vieira, E and J Gomes (2009), 'A comparison of Scopus and Web of Science for a typical university', *Scientometrics*, 81(2), pages 587-600.
Vieira, E S and J A N F Gomes (2010), 'Citations to scientific articles: its distribution and dependence on the article features', *Journal of Informetrics*, 4(1), pages 1-13.
von Eye, A and E Y Mun (2005), *Analyzing rater agreement. Manifest variable methods* (Lawrence Erlbaum Associates, Mahwah, NJ, USA).
Weller, A C (2002), *Editorial peer review: its strengths and weaknesses* (Information Today, Inc., Medford, NJ, USA).
Whitley, K M (2002), 'Analysis of SciFinder scholar and web of science citation searches', *Journal of the American Society for Information Science and Technology*, 53(14), pages 1210-1215.
Wilson, J D (1978), 'Peer review and publication', *Journal of Clinical Investigation*, 61(4), pages 1697-1701.

Table 1
Journals that published manuscripts that were submitted to but not published in ACP, by extent of changes (two coders) that the authors made to the manuscripts for publication elsewhere and Journal Impact Factors (JIF_3 2008 and JIF_5 2008)

Journal	Number of manuscripts	Extent of changes (median)*		Journal Impact Factor	
		Coder 1	Coder 2	JIF_3 2008	JIF_5 2008
Acta Chimica Slovenica	1	1	2	0.909	0.932
Acta Meteorologica Sinica	1	1	1	-[$]	-[$]
Advances in Atmospheric Sciences	1	1	1	0.679	0.921
Annales Geophysicae	2	1	1	1.660	1.758
Applied Optics	1	1	1	1.763	1.874
Atmospheric Environment	2	1.5	1	2.890	3.423
Atmospheric Research	2	1	1.5	1.456	1.466
Chemical Physics Letters	1	1	1	2.169	2.341
Chemosphere	1	1	1	3.054	3.445
China Particuology	1	1	1	-[$]	-[$]
Crystal Growth & Design	1	3	3	4.215	4.491
Current Nanoscience	1	3	3	2.437	2.760
Geophysical Research Letters	3	3	3	2.959	3.137
International Journal of Modern Physics B	2	1	1	0.558	0.473
Journal of Atmospheric and Solar-Terrestrial Physics	1	1	2	1.667	1.770
Journal of Environmental Engineering and Science	1	1	1	0.773	1.000
Journal of Environmental Monitoring	1	1	2	1.989	1.930
Journal of Geophysical Research	6	2	2.5	3.147	3.465
Journal of the Korean Meteorological Society	1	1	1	-[$]	-[$]
Meteorology and Atmospheric Physics	1	1	2	1.034	1.586
Molecular Physics	1	1	1	1.478	1.506
Science of the Total Environment	1	1	1	2.579	3.148
Tellus A	1	1	1	1.965	2.359
Tellus B	2	1	1	2.356	2.946
Theoretical and Applied Climatology	2	1	1.5	1.621	2.065
Total	38	1	1		

Notes.
[*] Codings for the assessment of the extent of changes: 1=the manuscript was published with no changes, with minimal (linguistic) changes, or with changes to the order of the sections that were predominately due to journal guidelines; 2=the manuscript was published with a medium extent of changes to the text, tables, and/or figures; 3=the content of the manuscript was published in connection with other research results.
According to guidelines for the interpretation of Kappa coefficients published by von Eye and Mun (2005), a coefficient of .56 indicates a good level of agreement of the two scientists' codings for the assessment of the extent of changes.
[$] The Journal Impact Factor is not available in the Journal Citation Reports from Thomson Reuters.

Table 2
Minimum (min), maximum (max), mean, standard deviation (sd), and median of citation counts for manuscripts published in ACP and ACPD (group 1), published in ACPD only or in ACPD and elsewhere (group 2), or published neither in ACP nor in ACPD but elsewhere (group 3). The citation counts were searched in the databases Chemical Abstracts (CA), Science Citation Index (SCI), Scopus, and Google Scholar (GS) for a fixed three-year citation window (3 years) and for the period from date of publication up to the end of 2008 (all years)

Group	Statistic	CA		SCI		Scopus		GS
		3 years	All years	3 years	All years	3 years	All years	All years
Group 1	n	958.00	958.00	958.00	958.00	951.00	958.00	932.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	max	167.00	282.00	176.00	290.00	207.00	340.00	361.00
	mean	8.49	13.36	9.72	15.41	11.87	18.13	22.86
	sd	11.32	17.57	12.99	19.51	15.68	22.59	28.56
	median	**6.00**	**9.00**	**6.00**	**10.00**	**7.00**	**12.00**	**16.00**
Group 2	n	74.00	74.00	74.00	74.00	51.00	52.00	52.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	max	12.00	14.00	12.00	17.00	17.00	26.00	27.00
	mean	1.76	2.30	2.04	2.76	3.82	5.23	6.46
	sd	2.69	3.58	2.83	3.91	4.47	5.98	5.63
	median	**1.00**	**1.00**	**1.00**	**1.00**	**2.00**	**3.00**	**5.50**
Group 3	n	17.00	34.00	17.00	34.00	15.00	32.00	31.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	max	8.00	29.00	9.00	25.00	10.00	17.00	22.00
	mean	1.29	2.74	1.71	3.32	2.73	2.38	4.90
	sd	2.20	5.66	2.37	5.11	2.74	3.65	4.74
	median	**0.00**	**1.00**	**1.00**	**1.50**	**2.00**	**1.00**	**3.00**
Total	N	1049.00	1066.00	1049.00	1066.00	1017.00	1042.00	1015.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	max	167.00	282.00	176.00	290.00	207.00	340.00	361.00
	mean	7.90	12.26	9.05	14.14	11.33	17.00	21.47
	sd	11.01	17.04	12.63	18.93	15.33	22.05	27.80
	median	5.00	8.00	6.00	9.00	7.00	11.00	14.00

Table 3
Description of the dependent and independent variables for negative binomial regression models with editorial decision as independent variable predicting Chemical Abstracts citations since publication (model A) and for a fixed three-year window (model B), Science Citation Index citations since publication (model C) and for a fixed three-year window (model D), Scopus citations since publication (model E) and for a fixed three-year window (model F), as well as Google Scholar citations since publication (model G)

Variable	Arithmetic mean or percent	Standard deviation	Minimum	Maximum
Model A: Chemical Abstracts citations since publication (outcome variable)				
Citations	13.02	17.43	0	282
Editorial decision	97%		0	1 (published in ACP)
Number of co-authors	6.04	5.21	1	66*
Number of pages	14.73	14.89	3	431*
CA section:				
Air pollution and industrial hygiene	47%		0	1
Mineralogical and geological chemistry	21%		0	1
Other section	2%		0	1
Unknown (reference category)$	30%		0	1
Model B: Chemical Abstracts citations for a fixed three-year window (outcome variable)				
Citations	8.36	11.26	0	167
Editorial decision	98%		0	1 (published in ACP)
Number of co-authors	6.07	5.23	1	66*
Number of pages	14.77	14.99	3	431*
CA section:				
Air pollution and industrial hygiene	47%		0	1
Mineralogical and geological chemistry	21%		0	1
Other section	2%		0	1
Unknown (reference category)$	30%		0	1
Model C: Science Citation Index citations since publication (outcome variable)				
Citations	15.02	19.35	0	290
Editorial decision	97%		0	1 (published in ACP)
Number of co-authors	6.04	5.21	1	66*
Number of pages	14.73	14.89	3	431*
CA section:				
Air pollution and industrial hygiene	47%		0	1
Mineralogical and geological chemistry	21%		0	1
Other section	2%		0	1
Unknown (reference category)$	30%		0	1
Model D: Science Citation Index citations for a fixed three-year window (outcome variable)				
Citations	9.57	12.92	0	176
Editorial decision	98%		0	1 (published in ACP)
Number of co-authors	6.07	5.23	1	66*
Number of pages	14.77	14.99	3	431*
CA section:				
Air pollution and industrial hygiene	47%		0	1
Mineralogical and geological chemistry	21%		0	1
Other section	2%		0	1
Unknown (reference category)$	30%		0	1

Continuation of Table 3
Description of the dependent and independent variables for negative binomial regression models with editorial decision as independent variable predicting Chemical Abstracts citations since publication (model A) and for a fixed three-year window (model B), Science Citation Index citations since publication (model C) and for a fixed three-year window (model D), Scopus citations since publication (model E) and for a fixed three-year window (model F), as well as Google Scholar citations since publication (model G)

Variable	Arithmetic mean or percent	Standard deviation	Minimum	Maximum
Model E: Scopus citations since publication (outcome variable)				
Citations	17.66	22.43	0	340
Editorial decision	97%		0	1 (published in ACP)
Number of co-authors	6.04	5.21	1	66*
Number of pages	14.75	14.91	3	431*
CA section:				
Air pollution and industrial hygiene	47%		0	1
Mineralogical and geological chemistry	21%		0	1
Other section	2%		0	1
Unknown (reference category)$	30%		0	1
Model F: Scopus citations for a fixed three-year window (outcome variable)				
Citations	11.73	15.60	0	207
Editorial decision	99%		0	1 (published in ACP)
Number of co-authors	6.09	5.25	1	66*
Number of pages	14.79	15.06	3	431*
CA section:				
Air pollution and industrial hygiene	47%		0	1
Mineralogical and geological chemistry	21%		0	1
Other section	2%		0	1
Unknown (reference category)$	30%		0	1
Model G: Google Scholar citations since publication (outcome variable)				
Citations	22.42	28.36	0	361
Editorial decision	98%		0	1 (published in ACP)
Number of co-authors	6.10	5.23	1	66*
Number of pages	14.84	15.12	4	431*
CA section:				
Air pollution and industrial hygiene	47%		0	1
Mineralogical and geological chemistry	21%		0	1
Other section	2%		0	1
Unknown (reference category)$	30%		0	1

Notes.
* These large numbers of co-authors and pages belong to special issues of ACP.
$ As for many manuscripts no section is given in Chemical Abstracts, the category "unknown" was used in the analysis (Marsh *et al*, 2009).

Table 4
Negative binomial regression models with editorial decision as independent variable predicting Chemical Abstracts citations since publication (model A) and for a fixed three-year window (model B), Science Citation Index citations since publication (model C) and for a fixed three-year window (model D), Scopus citations since publication (model E) and for a fixed three-year window (model F), as well as Google Scholar citations since publication (model G)

	Model A	Model B	Model C	Model D	Model E	Model F	Model G
Editorial decision (1=ACP)	1.394** (2.91)	1.627*** (4.60)	1.344*** (4.04)	1.499*** (5.89)	1.963*** (7.10)	1.205*** (3.93)	1.460*** (7.80)
Number of co-authors	0.0316*** (3.64)	0.0433*** (3.79)	0.0267** (3.27)	0.0393*** (3.48)	0.0269*** (3.57)	0.0376*** (4.79)	0.0281** (3.25)
Number of pages	0.0323** (3.06)	0.0334** (2.62)	0.0324** (3.12)	0.0352** (2.77)	0.0329*** (3.36)	0.0363** (2.89)	0.0323** (2.94)
CA section:							
Air pollution and industrial hygiene	0.563*** (5.99)	0.313*** (4.16)	0.268*** (3.36)	0.0106 (0.15)	0.205* (2.21)	-0.0511 (-0.73)	0.120 (1.52)
Mineralogical and geological chemistry	0.458** (3.23)	0.278** (2.89)	0.245* (2.16)	0.0711 (0.82)	0.114 (1.15)	-0.000248 (-0.00)	-0.0394 (-0.42)
Other section	0.446 (1.89)	0.208 (1.04)	0.266 (1.32)	0.0403 (0.22)	0.0924 (0.44)	-0.0764 (-0.48)	0.0293 (0.16)
Publication year	(exposure)		(exposure)		(exposure)		(exposure)
Intercept	-7.541*** (-16.30)	-0.554 (-1.54)	-7.108*** (-22.74)	-0.0956 (-0.38)	-7.508*** (-26.64)	0.442 (1.43)	-6.702*** (-29.21)
n	989	976	989	976	987	966	955

Note. ML-point estimates (the results of the *z*-test shown in parentheses).
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Interpretation example for the parameter estimates in the table: In model A the number of pages of a manuscript has a statistically significant effect on receiving citations with a parameter estimate of 0.0323. This means that for an additional page, the odds of receiving citations increase by a factor of 1.03 (=exp(0.0323)), holding all other variables in model A constant.

Table 5
Minimum (min), maximum (max), mean, standard deviation (sd), and median of citation counts for manuscripts with different median reviewers' ratings. The citation counts were searched in the databases Chemical Abstracts (CA), Science Citation Index (SCI), Scopus, as well as Google Scholar (GS) for a fixed three-year citation window (3 years) and for the period from date of publication up to the end of 2008 (all years)

Median rating[$]	Statistic	CA		SCI		Scopus		GS
		3 years	All years	3 years	All years	3 years	All years	All years
1	N	40.00	40.00	40.00	40.00	40.00	40.00	40.00
	min	1.00	1.00	1.00	1.00	1.00	1.00	3.00
	max	61.00	63.00	64.00	97.00	92.00	146.00	171.00
	mean	13.65	16.85	16.90	21.10	21.45	26.23	34.75
	sd	14.62	16.80	18.10	21.77	22.92	28.77	37.87
	median	**8.50**	**11.50**	**10.50**	**15.50**	**13.00**	**16.00**	**17.50**
1.5	N	57.00	57.00	57.00	57.00	57.00	57.00	57.00
	min	0.00	1.00	0.00	1.00	1.00	1.00	1.00
	max	95.00	99.00	102.00	102.00	122.00	122.00	203.00
	mean	13.18	15.25	15.23	17.25	18.89	20.35	26.98
	sd	16.64	19.23	18.85	20.48	23.61	23.67	33.03
	median	**7.00**	**9.00**	**11.00**	**13.00**	**12.00**	**14.00**	**19.00**
2	N	243.00	243.00	243.00	243.00	243.00	243.00	234.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	max	92.00	142.00	146.00	224.00	179.00	265.00	361.00
	mean	9.13	11.10	10.56	13.01	13.21	15.70	21.42
	sd	10.77	13.74	13.59	18.05	15.64	20.49	31.12
	median	**7.00**	**8.00**	**7.00**	**8.00**	**9.00**	**11.00**	**15.00**
2.5	N	86.00	86.00	86.00	86.00	86.00	86.00	84.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	max	56.00	56.00	74.00	82.00	88.00	94.00	127.00
	mean	8.56	9.50	9.64	10.98	12.17	13.49	17.25
	sd	9.78	10.56	11.93	12.91	14.79	15.81	20.76
	median	**5.50**	**7.50**	**5.50**	**7.00**	**7.00**	**9.00**	**11.50**
3	N	70.00	70.00	70.00	70.00	70.00	70.00	67.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	max	46.00	59.00	53.00	67.00	85.00	87.00	103.00
	mean	7.13	8.94	7.70	9.89	10.24	12.26	16.48
	sd	7.92	9.85	8.11	10.18	11.66	12.91	16.06
	median	**5.00**	**5.00**	**6.00**	**7.00**	**8.00**	**9.00**	**14.00**
Total	N	496.00	496.00	496.00	496.00	496.00	496.00	482.00
	min	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	max	95.00	142.00	146.00	224.00	179.00	265.00	361.00
	mean	9.58	11.46	11.05	13.35	13.93	16.21	21.77
	sd	11.58	13.95	14.03	17.22	17.05	20.34	29.10
	median	6.00	8.00	7.00	9.00	9.00	11.00	15.00

Note.
[$] 1=excellent, 2=good, 3=fair to poor. As for only 14 manuscripts the median was greater than 3, all manuscripts with a median of at least 3 were grouped in one category.

Table 6
Description of the dependent and independent variables for negative binomial regression models with reviewers' median ratings as independent variable predicting Chemical Abstracts citations since publication (model A) and for a fixed three-year window (model B), Science Citation Index citations since publication (model C) and for a fixed three-year window (model D), Scopus citations since publication (model E) and for a fixed three-year window (model F), as well as Google Scholar citations since publication (model G)

Variable	Arithmetic mean or percent	Standard deviation	Minimum	Maximum
Model A: Chemical Abstracts citations since publication (outcome variable)				
Citations	11.46	13.95	0	142
Model B: Chemical Abstracts citations for a fixed three-year window (outcome variable)				
Citations	9.58	11.58	0	95
Model C: Science Citation Index citations since publication (outcome variable)				
Citations	13.36	17.22	0	224
Model D: Science Citation Index citations for a fixed three-year window (outcome variable)				
Citations	11.05	14.03	0	146
Model E: Scopus citations since publication (outcome variable)				
Citations	16.21	20.34	0	265
Model F: Scopus citations for a fixed three-year window (outcome variable)				
Citations	13.93	17.05	0	179
Independent variables for models A to F				
Reviewers' ratings	2.11	0.58	1 (excellent)	4 (poor)
Number of co-authors	6.68	6.04	1	66*
Number of pages	15.79	19.94	4	431*
CA section:				
Air pollution and industrial hygiene	40%		0	1
Mineralogical and geological chemistry	22%		0	1
Other section	2%		0	1
Unknown (reference category)$	36%		0	1
Model G: Google Scholar citations since publication (outcome variable)				
Citations	21.77	29.10	1	361
Independent variables for model G				
Reviewers' ratings	2.11	0.59	1 (excellent)	4 (poor)
Number of co-authors	6.71	6.05	1	66*
Number of pages	15.91	20.20	4	431*
CA section:				
Air pollution and industrial hygiene	41%		0	1
Mineralogical and geological chemistry	22%		0	1
Other section	2%		0	1
Unknown (reference category)$	35%		0	1

Notes.
* These large numbers of co-authors and pages belong to special issues of ACP.
$ As for many manuscripts no section is given in Chemical Abstracts, the category "unknown" was used in the analysis (Marsh *et al*, 2009).

Table 7
Negative binomial regression models with reviewers' median ratings as independent variable predicting Chemical Abstracts citations since publication (model A) and for a fixed three-year window (model B), Science Citation Index citations since publication (model C) and for a fixed three-year window (model D), Scopus citations since publication (model E) and for a fixed three-year window (model F), as well as Google Scholar citations since publication (model G)

	Model A	Model B	Model C	Model D	Model E	Model F	Model G
Reviewers' ratings	-0.255***	-0.242***	-0.274***	-0.278***	-0.274***	-0.268***	-0.266***
	(-3.53)	(-3.40)	(-3.64)	(-3.72)	(-3.45)	(-3.52)	(-3.94)
Number of co-authors	0.0255*	0.0321**	0.0219	0.0289*	0.0185	0.0246*	0.0211*
	(2.15)	(2.69)	(1.81)	(2.39)	(1.56)	(2.16)	(1.99)
Number of pages	0.0184	0.0183	0.0204	0.0208	0.0207	0.0221	0.0197
	(1.01)	(1.01)	(1.12)	(1.13)	(1.13)	(1.21)	(1.47)
CA section:							
Air pollution and industrial hygiene	0.385**	0.295*	0.136	0.0151	0.154	0.0236	0.0392
	(2.88)	(2.32)	(1.09)	(0.12)	(1.37)	(0.22)	(0.32)
Mineralogical and geological chemistry	0.133	0.111	-0.0412	-0.0581	0.00107	-0.0454	-0.208
	(0.89)	(0.71)	(-0.28)	(-0.37)	(0.01)	(-0.32)	(-1.20)
Other section	0.0214	0.0816	-0.341	-0.276	-0.179	-0.166	-0.531
	(0.09)	(0.34)	(-1.36)	(-1.16)	(-0.61)	(-0.64)	(-1.09)
Publication year	(exposure)		(exposure)		(exposure)		(exposure)
Intercept	-5.344***	2.047***	-5.008***	2.403***	-4.808***	2.624***	-4.446***
	(-16.96)	(6.44)	(-17.31)	(8.05)	(-16.16)	(8.67)	(-15.73)
n	496	496	496	496	496	496	482

Note. ML-point estimates (the results of the *z*-test in parentheses).
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Interpretation example for the parameter estimates in the table: In model A the number of co-authors for a paper has a statistically significant effect on receiving citations with a parameter estimate of 0.0255. This means that for an additional co-author, the odds of receiving citations increase by a factor of 1.03 (=exp(0.0255), holding all other variables in model A constant.