Code Sharing in Psychological Methods and Statistics: An Overview and Associations with Conventional and Alternative Research Metrics

Šimon Kucharský*, Bobby-Lee Houtkoop, and Ingmar Visser

Department of Psychology, University of Amsterdam, Amsterdam

*Correspondence concerning this article may be addressed to Šimon Kucharský, University of Amsterdam, Department of Psychology, Nieuwe Achtergracht 129-B, 1018 WS, the Netherlands. E-mail may be sent to s.kucharsky@uva.nl

Abstract

Towards discovering more effective means to promote code sharing as an open science practice, this study explores the current state of code sharing within the field of psychological methods and statistics and examines the association between this practice and conventional and alternative research metrics. Towards this purpose, a total of 815 articles from three major journals within the field of psychological methods and statistics were manually screened and encoded based on code sharing practices and general article characteristics. In addition, data on conventional (citation counts) and alternative research metrics (Altmetric Attention Score) was retrieved through online scientific databases (Web of Science and Altmetric.com). This input was then analysed using descriptive statistics and regression models suitable for count data, and robustness of the findings were assessed using multiverse analysis.

The findings of this study suggest that the sharing of scientific computer code is not (yet) extensively practiced within the field of psychological methods and statistics. In the majority of academic articles included in this study, scientific computer code was not shared (66 %). Moreover, if such code was shared, it was frequently found to be improperly annotated (70 %) and/or incomplete (52 %). Nevertheless, the findings of this study also suggest a hopeful prospect, as the sharing of scientific computer code has increased between 2010 and 2017. The study revealed a robust positive connection with alternative research metrics. This study did not find robust positive connection between code sharing and citation counts.

Keywords: Open Science Practices, Code Sharing, Open Science, Research Metrics, Citation Rate, Alternative Research Metrics, Altmertic

1 Introduction

The open science movement is gaining momentum. The adoption of open science practices by journals, funders, and institutions in psychology is growing, and an increasing number of studies focus on promoting a more transparent way of practicing science in a diverse range of psychological disciplines (McKiernan et al., 2016). Prominent themes within the open science movement are pre-registration and the sharing of primary research data. As more disciplines tend to depend more heavily on computer code for, among other things, simulations, statistical analyses, and data visualization, the need for open practices regarding the sharing of scientific computer code is growing as well (Opensource.com, n.d.).

1.1 Code Sharing

In this paper, code sharing is described as the practice of the sharing of computer code that a researcher has developed in any stage during the research process. This computer code can be written in any programming language, can be of any level of complexity, and may include code dependencies (i.e., the code depends on external software packages; Chawla, 2016).

Sharing of such scientific computer code could constitute an effective means to promote the reproducibility of research articles in psychology, and decrease the risk of errors within the code and thus in the results of the analyses (Barnes, 2010; Challet & Du, 2003; Epskamp, 2019; Veldkamp, Nuijten, Dominguez-Alvarez, van Assen, & Wicherts, 2014). Moreover, code writing is an important part scientific endeavor that may require large effort. Sharing of scientific computer code makes this work more visible, allowing individual researchers that wrote the code to gain recognition for their contribution to science.

Despite the advantages of code sharing in terms of reproducibility, scientific computer code is still rarely shared (Shamir et al., 2013). To change this practice, a number of initiatives is concerned with incentivizing the sharing of computer code, for example through researcher led initiatives and journal policies.

A researcher led initiative that focuses on directly rewarding researchers who share their computer code is the Badges Project (Blohowiak et al., 2013; Kidwell et al., 2016). The Badges project is a means for journals to encourage open science practices through actively acknowledging such practices, using badge symbols which are added at the top of an authors' article. These badges are divided into three types: the Open Data badge, the Open Materials badge and the Preregistered badge (Blohowiak et al., 2013). Code sharing can be rewarded with the Open Materials badge. In terms of effectiveness, Kidwell et al. (2016, p. 10) found that the use of badges in journals "increased the rate of actual, correct, usable, and complete sharing of materials compared to journals without badges".

A second initiative that aims to support journals in the promotion of code sharing is the creation of the

TOP guidelines (<https://cos.io/top/>). The TOP guidelines consist of eight standards that journals can adopt to varying extent to “move scientific communication toward greater openness” (Nosek et al., 2015, p. 1423). One of the eight transparency standards relates to the encouragement of code sharing; the analytic methods standards. These analytic methods standards aim to “recognize the value of the availability of the code comprising the statistical models or simulations conducted for the research” (Nosek et al., 2015, p. 1423).

The initiatives that attempt to directly motivate researchers for adopting open science practices have not necessarily led to a noticeable increase in code sharing (Shamir et al., 2013). A potential cause of this is that researchers experience barriers in the practice of sharing scientific computer code. Such barriers for code sharing are: (1) sharing is not a common practice, (2) people will pick holes and demand support and bug fixes, (3) the code is valuable intellectual property that belongs to my institution, and (4) it is too much work to polish the code (Barnes, 2010; LeVeque, 2013). Further, sharing a computer code that is reproducible for a long period of time and across platforms might require additional work and expertise (Epskamp, 2019).

1.2 Impact of Code Sharing

To overcome such barriers, researchers need to be motivated through incentives directly aimed at the potential benefits they could obtain from sharing code, relative to the increased effort associated with this practice. Important motivators in the current scientific system are research metrics such as citation rate, media attention or funding opportunities. However, up to this moment, no specific research on the effects of code sharing on research metrics has been conducted. Other open practices have been related to the improvement of research metrics in a small number of studies, with the goal to motivate researchers to adopt transparent practices more frequently.

Prompted by the increasing relevance of the internet as a medium for publication of scientific articles, Lawrence (2001) investigated the impact of free online availability of research articles on citation rates of a set of 119,924 conference articles in computer science. In this study, Lawrence (2001) found that the mean number of citations for offline articles was 2.74 and the mean number of citations for online articles is 7.03. Likewise, Piwowar, Day, and Fridsma (2007) found that, independent from the impact factor of the journal, the date of publication, and author country of origin, publicly available data was significantly associated with a 69% increase in citation counts.

More recent work by McKiernan et al. (2016) reviewed literature that studied the effects of various open research practices on research metrics such as citations, media attention, potential collaborators, job opportunities, and funding opportunities. In accordance with studies by Lawrence (2001) and Piwowar et

al. (2007), McKiernan et al. (2016) reported several studies that found a relationship between open access publishing and citation rates (Hitchcock, 2016; Adie, 2014; Eysenbach, 2006; Hajjem et al., 2006; cited in McKiernan, 2016). Moreover, McKiernan et al. (2016) found open research practices to be associated with alternative research metrics as well. For example, open publications get more media coverage than articles that are not openly accessible (Adie, 2014a; Wang et al., 2015; cited in McKiernan et al., 2016).

Lastly, to our knowledge, the only study that focuses on the specific effects of code sharing was conducted by Vandewalle (2012). In this study, Vandewalle performed two studies on the effects of code sharing on citation rate for articles in signal and image processing. In the first study, code sharing was found to be positively related with citation rate. In the second study, Vandewalle found that 87 % of the best-cited image processing articles had code available online. These results suggest that “studies with openly published code are also more likely to be cited than those that do not open their code” (Vandewalle, 2012).

1.3 The Current Study

Although the above mentioned findings are mostly observational, they may suggest that compared to traditional (“closed”) practices, open research practices can in general bring significant benefit to researchers by affecting both conventional research metrics (such as citation rates), and alternative research metrics (such as media coverage). Moreover, they also show that an additional research regarding code sharing practices in psychology is needed. Therefore, this article presents an overview of the current prevalence of code sharing in (a subset of) psychological literature, and focuses on determining the association of sharing scientific code with conventional and alternative research metrics.

Towards this purpose, we formulate the following research questions: (1) “What is the current status of code sharing practices in the field of psychological methods and statistics?”, and (2) “What is the association of sharing scientific computer code with both conventional research metrics and alternative research metrics in the field of psychological methods and statistics?”.

To answer these questions, we analyzed a set of 815 articles from three major journals within the field of psychological methods and statistics regarding code sharing practices. Further, we related the code sharing practices to the measures of conventional and alternative research metrics that were retrieved through the Web of Science (<https://apps.whoofknowledge.com/>) and Altmetric (www.altmetric.com) databases, respectively.

2 Methods

2.1 Data Collection Procedure

The current study focuses on the collection and analysis of four sets of variables; (1) code sharing related variables, (2) conventional research metrics, (3) alternative research metrics, and (4) general article characteristics (see section 2.2 for details). As a first step, access was acquired to the online database Altmetric that collects alternative research metrics variables for a wide variety of articles. Altmetric monitors a range of non-traditional sources (such as mainstream media, blog posts, and policy documents) by searching for links and references to published research (Altmetric, n.d.-b).

For this study, Altmetric data was retrieved for articles from three journals within the field of psychological methods and statistics: *Psychological Methods*, *Behavior Research Methods*, and *The American Statistician* in a timespan from 01/01/2010 – 31/12/2017. These choices as search settings in the Altmetric database resulted in a total set of 960 articles (search executed on February 26, 2018). It is important to note that the Altmetric database does not record articles of which all Altmetric variables have a value of zero, meaning they were not yet mentioned in any non-traditional source that Altmetric is tracking. However, articles of Altmetric customers constitute an exception to this rule. Even if these articles have values of zero for each variable, Altmetric includes them in the database to ensure customers their articles are being tracked (A. Rees, personal communication, April 17, 2018).

Consequently, whether or not the Altmetric database includes data for an article depends on the following conditions: (1) the article is published in a journal that is tracked by Altmetric, and (2) the article is mentioned in at least one tracked alternative source or the article is tracked for a paying customer.

After the alternative research metrics data (e.g. mentions on social media) and general article characteristics (e.g. publication year) were retrieved from Altmetric, and the Digital Object Identifiers (DOIs) were used to retrieve the citation counts from Web of Science. For three articles, the DOI was unknown, which is why they were removed from the dataset. From the 957 articles that did have a known DOI, 19 articles had missing citation rates. These articles were excluded from the dataset as well, resulting in a total of 938 articles. From these 938 articles, a random sample of 100 articles was used for the pilot study. Additionally, a total of 23 articles were excluded from the data set as Web of Knowledge had no records for these articles, resulting in a final sample of 815 articles which were included in the main study.

These 815 articles were manually screened to collect data for the set of code sharing variables and general article characteristics such as the type of article. The screening of these articles was done following four general steps: (1) by reading the abstract of the article to determine the type of article, (2) by checking whether or not the article provides supplemental materials, (3) by using control + F to search for specific

key words¹ which were selected during the pilot study, and (4) by scrolling through the entire article to see if any shared code was missed during the first three steps.

The sample of the main study consists of research articles. The total set of 815 articles are all published in one of three selected journals in the field of psychological methods: (1) 558 articles from Behavior Research Methods (published by Springer), 128 articles from Psychological methods (published by APA), and (3) 129 articles from The American Statistician (published by Taylor & Francis Group).

The selected journals meet four predefined criteria: (1) the journals are concerned with methods and methodological innovations in psychology and/or statistics, (2) the journals publish, besides regular studies, tutorial articles related to certain software for statistical data-analysis, (3) the journals publish articles that are mostly targeted at larger audiences of, among others, applied researchers, and (4) all three journals state to “encourage” or “request” relevant computer code to be shared, but do not oblige authors to share (see Appendix 1 for the exact policies).

For all three journals included in this study, a time span was selected between 01/01/2010 – 01/01/2017. This time span restricts the number of articles that have to be coded, while also ensuring the studied samples provides a good representation of recently published articles.

2.2 Variables

This section entails a detailed description of the variables that data was retrieved for, either through the screening and coding of research articles, or the use of online scientific databases.

Code Sharing. All code sharing variables were coded by the researcher through the screening of research articles. The categories for the coding were determined during a pilot study with a set of 100 articles, that followed the same procedures as the described main study.

The first variable is the occurrence of code sharing (1 = “Yes, and it is accessible”, 2 = “Yes, but it is not accessible (anymore)”, 3 = “No, and there is no code to be shared”, 4 = “No, but there was code to be shared”, and 5 = “Code available upon request”). Next, the following variables were assessed: (1) the type of code (1 = “NA” (this value is assigned to all code characteristic variables if code is not shared or there was no code to be shared), 2 = “Code is not accessible (CNA)”, 3 = “Data analysis script”, 4 = “Software”, 5 = “Simulation and analysis”, 6 = “Tutorial Code”, 7 = “Multiple”, 8 = “Other”), (2) the programming language², (3) the complexity of the code (1 = “NA”, 2 = “CNA”, 3 = “Low”, 4 = “High”), and (4) the

¹“Code”, “data”, “syntax”, “shared”, “available”, “open source”, “transparency”, “online appendix”, “online supplemental material”, “appendix”, “supplemental materials”, “download”, “website”, “be found”, “archive”, “package”, “www.”, “http”, “script”, “analyses”, “computed”, “executed”.

²1 = NA, 2 = CNA (code is not accessible), 3 = R, 4 = MATLAB, 5 = Python, 6 = SAS, 7 = MPLUS, 8 = STATA, 9 = JASP, 10 = RUBY, 11 = JAVA, 12 = C, 13 = CC+, 14 = Other, 15 = Stan, 16 = Multiple languages, 17 = Java, 18 = SPSS.

way in which code is shared (1 = “NA”, 2 = “CNA”, 3 = “Website of institution”, 4 = “Personal website”, 5 = “Online supplemental material”, 6 = “Appendix of article”, 7 = “In text”, 8 = “Online repository”, 9 = “Website dedicated to shared code”, 10 = “Website other”, 11 = “Multiple”), (5) whether or not the provided code is complete (1 = “NA”, 2 = “CNA”, 3 = “Yes”, 4 = “No”), (6) whether or not the code is annotated (1 = “NA”, 2 = “CNA”, 3 = “None”, 4 = “Some”, 5 = “A lot”).

Conventional Research Metrics. The conventional research metric used in this study is citation counts. The citation count relates to the number of times an article is cited in another academic articles. Citation counts were retrieved through Web of Science, using the DOIs of the articles (on February 26, 2018).

Alternative Research Metrics. Altmetric records count data for the following variables: the number of (1) mentions in news stories in mainstream media, (2) mentions in academic and non-academic blog posts, (3) mentions in policy documents, (4) peer reviews on post-publication peer-review platforms (www.pubpeer.com and www.publons.com), (5) mentions in social media (Facebook (mentions on public pages only), Twitter, Google+, LinkedIn, Weibo, and Pinterest), (6) mentions in multimedia (YouTube, Reddit, and StackOverflow), (7) mentions in English Wikipedia pages, (8) mentions in syllabi of institutions (from www.opensyllabusproject.org), (9) research highlights (number of recommendations on www.F1000.com) and (1) readers in online reference managers (Mendeley.com) (Altmetric, n.d.-a). Ultimately, Altmetric uses this data to calculate the “Altmetric Attention Score”, which is “a weighted count of the amount of attention we’ve picked up for a research output” (Altmetric, 2016). The Altmetric Attention Score (AAS) constitutes the second outcome variable in this study. Due to Altmetric restriction policy, we are allowed to share only the AAS from the data set acquired through the Altmetric database (i.e., the individual variables collected by Altmetric can be accessed only by customers of Altmetric).

General Article Characteristics. Lastly, we collected the general characteristics of the articles. These general characteristics were documented in order to provide relevant contextual information on the current status of code sharing.

The Altmetric data is used to retrieve data for the following variables: the author(s), DOI, journal and the publication year of the articles. Additionally, the type of article was manually recorded (1 = “Empirical article”, 2 = “General nonempirical article”, 3 = “Tutorial”, 4 = “Review”, 5 = “Commentary”, 6 = “Editorial”, 7 = “Perspective/Opinion”). These categories were based on the article types described in the submission guidelines of the journal *Advances in Methods and Practices in Psychological Science* (Association for Psychological Science, 2016).

Additionally, various open science practices were manually recorded; data sharing (0 = “No”, 1 = “Yes”), material sharing (0 = “No”, 1 = “Yes”), and the sharing of any other files, such as extra tables or theoretical background (0 = “No”, 1 = “Yes”). It is important that any shared data or materials are primary documents; the sharing of secondary documents (a document that was created by another author) does not qualify as the sharing of data or materials.

Lastly, all articles were coded on whether or not they provided proper citations for mentioned or used software (1 = “No software used/mentioned”, 2 = “No”, 3 = “Yes”, 4 = “Yes, for some used/mentioned software, but not for all”). In relation to these citations, the APA publication manual provides the following general statement: “Reference entries are not necessary for standard software and programming languages, such as Microsoft Word or Excel, Java, Adobe Photoshop, and even SAS and SPSS”. Following this guideline, Microsoft Word, Excel, and Adobe Photoshop were categorized in this study as standard software, while all other software was categorized as being specialized. All specialized software was checked for proper citations. The value “Yes” was assigned when the used and/or mentioned software was cited in text, as well as in the reference list.

2.3 Data Analytic Approach

Descriptive Analyses. To investigate the current status of the open science culture regarding code sharing, descriptive analyses were used. This includes bar plots with regard to the following variables: (1) open science practices in general, (2) code sharing practices, (3) characteristics of code sharing (e.g. programming language, type of code, and annotation), (4) references to used or mentioned software. Additionally, histograms of the variables citation counts and Altmetric Attention Score were included to explore their distributions and identify possible outliers.

Hypothesis Testing. To study the effects of code sharing on conventional and alternative research metrics, hypotheses 1 and 2 were tested.

H1: The sharing of scientific computer code in an article is associated with a higher citation rate of that article.

H2: The sharing of scientific computer code in an article is associated with a higher Altmetric Attention Score of that article.

Hypotheses 1 and 2 were tested with separate regression models. In these two models, code sharing was the categorical predictor variable, and the outcome variable was either citation counts or Altmetric Attention Score (AAS). To account for that older articles have higher citation counts and possibly AAS, both models

included the covariate publication year.³

The variables of interest are count variables, and hence linear regression models are not suitable. Instead, we used the Poisson and Negative Binomial likelihood to model the distribution of the dependent variable, and the log-link to specify the regression model on the parameters of the likelihood functions. The Poisson distribution describes the probability of a given number of events that occurred in a fixed time interval, under the assumption that these events occur independently and with a specific rate. However, the Poisson distribution implies that given a specific rate, the mean number of observed events equals the variance of the observed events. Given the data in the pilot study showed that this is not the case (i.e., the citation counts and AAS are “overdispersed”), we also use the Negative Binomial likelihood to model the overdispersion.

In addition, count data is sensitive to zero-inflation, which arises when the number of observed zeros is higher than expected under either of the Poisson or Negative binomial model (Hu, Pavlicova, & Nunes, n.d.). In the pilot data, both the CR and AAS variables showed zero-inflation, leading to an expectation of zero-inflation in the main data as well.

To account for zero-inflation in count data, we also used the zero-inflated models and hurdle models (Hu et al., n.d.). Zero-inflated models assume that observations of zero events originate from two sources: the structural source and the sampling source. Sampling zeros happen by chance and according to the Poisson or Negative binomial distribution, whereas structural zeros occur because zero is the only value that the data point can take. In contrast to zero-inflation models (which assume zero observations originate from both structural and sampling sources), hurdle models assume that all observed zeros come from a single source: the structural source (Hu et al., n.d.). This part of the data is usually modelled with a Bernoulli variable. The positive counts are modelled using a Poisson or negative binomial distribution truncated at 0 (to disable the option of observing 0 as a result of the sampling source).

In order to find the regression model that performs best, six models were considered: (1) Poisson regression model, (2) negative binomial regression model (NB), (3) zero-inflated Poisson model (ZIP), (4) zero-inflated negative binomial model (ZINB), (5) Poisson hurdle model (PH), and (6) negative binomial hurdle model (NBH).

The performance of these models was evaluated using the Bayesian Information Criterium (BIC) and Akaike Information Criterium (AIC, Akaike, 1998), and the results of AIC and BIC corrected Vuong tests (Vuong, 1989). The Vuong test tests the null hypothesis that two models perform equally well, against the alternative hypothesis that one model performs better (Vuong, 1989).⁴

³In all models, we rescaled the publication year by subtracting 2009 from each value, such that the intercept corresponds to the dependent variable in the year 2009.

⁴In the pilot study, the negative binomial hurdle model performed best for both the Altmetric Attention Score and citation rate models. Based on these pilot results, we expected that in the main study the negative binomial hurdle model will perform best for both outcome variables.

The two main hypotheses about association of code sharing with either citation counts or AAS were evaluated with p -values of the predictor in the best model. The criterion for statistical significance was set to $\alpha = 0.05$.

In addition to model evaluation and selection, we report the results of the hypothesis tests based on all six models for both dependent variables. To further assess the robustness of the findings against the multiverse of possible data-analytic choices, we also report the results with different criteria for excluding outliers. These outliers were identified by calculating the standardized Pearson residuals (under the model fit on the whole dataset), excluding observations with the absolute value of the residual above 2, 3, 5, or 10, and refitting the model. This results in reporting the total of 6 (model choice) $\times$ 5 (full data and 4 exclusion criteria) = 30 hypothesis tests per each dependent variable to assess the robustness of the main conclusions.

The analyses were conducted in R language (R Core Team, 2018) in the environment R Studio (RStudio Team, 2018). We used the following R packages: `tidyverse` packages (most notably `ggplot2`, `dplyr`, `tidyr`, `magrittr`, `purrr`, `readxl`, and `forcats`; Wickham, 2017), `broom` (Robinson & Hayes, 2019), `MASS` (Venables & Ripley, 2002), `pscl` (Jackman, 2017; Zeileis, Kleiber, & Jackman, 2008a), `yhat` (Nimon, Oswald, & Roberts., 2013), `vcdExtra` (Friendly, 2017), `countreg` (Zeileis & Kleiber, 2018; Zeileis et al., 2008a; Zeileis, Kleiber, & Jackman, 2008b), base R packages (`grid`, `stats`, `graphics`, `grDevices`, `utils`, `methods`, and `base`; R Core Team, 2018), `gridExtra` (Auguie, 2017), `knitr` (Xie, 2014, 2015, 2019), `kableExtra` (Zhu, 2019), and `rlang` (Henry & Wickham, 2019). For specific versions of the packages, see the R session info at <https://osf.io/cvhpz/>. All code that was written to produce the reported results and figures is openly accessible at the Open Science Framework (<https://osf.io/4keaq/>).

3 Results

3.1 Characteristics of Articles

The sample of 815 articles included one article that was published in 2008, which falls outside of the selected time span from 2010 to 2017. This article was removed from the dataset. The resulting sample of 814 articles included 322 empirical articles, 447 general non-empirical articles, 19 tutorials, 14 commentaries, eight reviews, two editorials, and two perspective/opinion articles. 557 articles were published in Behavior Research Methods (BRM), 128 articles in Psychological Methods (PM), and 129 articles in The American Statistician (TAS).

The majority of articles was published in BRM. This overrepresentation has two possible explanations. The first possible explanation is that BRM simply publishes more articles than PM and TAS, resulting in

the inclusion of more articles from this journal. The second possible explanation was provided by Altmetric, and pertains to the possibility that Altmetric has a larger number of paying customers for BRM than for PM and TAS (A. Rees, personal communication, April 17, 2018). The Altmetric database tracks articles that are mentioned in alternative research metrics sources, as well as articles from paying customers. Articles with an AAS of zero are only included in the Altmetric database if customers are paying to view the tracking results. Consequently, if BRM has more paying customers than the other two journals included in this study, this may have led to a larger number of included articles from this journal.

3.2 Descriptive Analyses

The data for the various open science practices, code sharing, data sharing, material sharing and the sharing of additional documents (“other”), suggest that open science practices remain uncommon. Out of 814 articles, code was shared in 244 instances (30 %), data was shared in 90 instances (11 %), materials were shared in 129 instances (16 %), and other files were shared in 250 instances (31 %) (Figure 1).

Figure 1: Frequencies of sharing practices: (a) code sharing, (b) material sharing, (c) data sharing, and (d) sharing of additional types of documents. NA encodes articles that did not produce anything to be shared.

Code sharing. The main study showed the following results for the five categories of code sharing: (1) 458 (56 %) articles in which the author(s) did not share code, despite the article containing code that could have been shared (for example, scripts for data analyses or figures), (2) 219 (27 %) articles which contained shared code that was accessible, (3) 89 (11 %) articles in which the author(s) did not use any computer code, (4) 25 (3 %) articles which contained shared code that was no longer accessible (e.g. due to broken URLs

or expired web hosting), and (5) 23 (3 %) articles which contained a statement indicating that the code was available upon request (Figure 2).

Figure 2: Frequencies for five different categories of the code sharing variable.

In further analyses, category 2 and 4 constitute the group “code is shared” ($N = 244$). Category 1 and 5 constitute the group “code is not shared” ($N = 481$). Articles that did not employ any form of computer code and thus could not share code (category 3), will be disregarded in further analyses (resulting in total of 725 articles that relied on some code).

Table 1 shows that about one third of the articles that relied on code of some sorts shared it, suggesting that code sharing is not yet a common open science practice within the field of psychological methods and statistics. However, Table 1 does show an increasing trend of code sharing practices, with a peak in 2016. A logistic regression analysis that models the log-odds of code sharing, including publication year as the only predictor, shown that this trend is statistically significant at the $\alpha = 0.05$ level; $\beta_{\text{year}} = 0.083$, 95% CI [0.012, 0.154], $p = 0.022$.

Table 1: Number of articles that shared and did not share code, the total number of articles and the percentage of articles that did share code for each included publication year.

Publication year	Nr. of articles using code	Code is shared	Code is not shared
2010	40	10 (25 %)	30 (75 %)
2011	81	22 (27 %)	59 (73 %)
2012	78	24 (31 %)	54 (69 %)
2013	91	29 (32 %)	62 (68 %)
2014	81	26 (32 %)	55 (68 %)
2015	90	32 (36 %)	58 (64 %)
2016	136	55 (40 %)	81 (60 %)
2017	128	46 (36 %)	82 (64 %)
<i>Total</i>	725	244 (34 %)	481 (66 %)

Figure 3 shows the code sharing over the time period of 2010 to 2017 for the three journals separately. Figure 3 shows the total counts and the proportion of code sharing. The figure suggests that the positive trend of increasingly popular code sharing is perhaps mainly driven by Behavior Research Methods (small increase in code sharing prevalence and a large total number of articles) and Psychological Methods (large increase in code sharing although relatively few articles) journals, although we did not test this pattern

statistically.

Figure 3: Frequencies of articles that shared code with publication year on the x-axis, displayed for each analysed journal separately. The figure in the left panel shows the counts of articles, the figure on right shows the proportions within each year.

Of the total number of 244 articles that shared code in some format, the most common programming languages were R (87; 36 %) and MATLAB (34; 14 %) (Figure 4a). Shared code consisted in most cases of data analysis scripts (102; 42 %), followed by open source software code (70; 29 %), and simulation scripts (27; 11 %) (Figure 4b). In terms of location, code was most frequently shared via online supplemental materials (73; 30 %), online repositories (54; 22 %) or the appendix of the article (44; 18 %) (Figure 4c). The complexity of the shared code was categorized as “high” in the majority of cases (207; 85 %) (Figure 4d). Annotation of code typically consisted of some notes, but often not enough to fully understand the entire code (109; 45 %). Moreover, 25 % (61) of the shared code was not annotated at all (Figure 4e). Lastly, in the majority of cases code was only shared partially, with either a part of the analyses or the code to produce the figures missing (128; 52 %) (Figure 4f).

The entire data set ($N = 814$) was used to examine whether or not the use of existing code or software was properly referenced. Software was not referenced in 415 (51 %) articles, referenced only partially in 187 (23 %) articles, and properly referenced in 154 (19 %) articles. The remaining 58 articles (7 %) did not use or mention any software.

Citation counts and Altmetric Attention Score. Next, the dependent variables citation counts and Altmetric Attention Score were explored using histograms. Both variables show extreme outliers and a high number of zero observations.

We used the `boxplot` function from `grDevices` package to identify outliers separately for the citation counts and AAS. These outliers were removed prior the main analyses. 79 extreme observations of citation counts were removed prior analyzing the citation counts, resulting in a data set of 646 observations. 101 extreme AAS observations were removed prior analyzing the AAS, resulting in a data set of 624 observations.

Figure 4: Six characteristics of shared code: (a) programming language or program, (b) the type of code, (c) the location of the code, (d) the complexity of code, (e) whether or not the code is complete, and (f) the extent to which code is annotated.

Table 2 shows that in both variables Altmetric Attention Score (AAS) and citation counts (CC), the variance exceeds the mean. This indicates that the data for the variables AAS and CC is overdispersed.

Table 2: Descriptive statistics for the variables citation rate and AAS, for articles that did and did not share computer code.

	Code Sharing	Mean	SD	Variance
Citation Rate <i>N</i> = 646	No	6.44	7.81	60.96
	Yes	6.39	7.17	51.44
	Total	6.42	7.59	57.68
AAS <i>N</i> = 624	No	1.53	1.34	1.80
	Yes	1.82	1.47	2.16
	Total	1.62	1.39	1.93

3.3 Association of code sharing with citation counts

Model Selection. As stated in the methods section, we used six different regression models for count data: (1) Poisson regression (P), (2) negative binomial regression (NB), (3) zero-inflated Poisson regression (ZIP), (4) zero-inflated negative binomial regression (ZINB), (5) Poisson hurdle regression (PH), and (6) negative binomial hurdle regression (NBH). We compared the models using AIC, BIC, and a Vuong test. Further, we assessed the adequacy of the best performing model using rootograms.

The ZINB model and the NBH model show the lowest AIC and BIC values (Table 3). The Vuong test

Table 3: Comparison of the six models: P - Poisson, NB - Negative binomial, ZIP - Zero-inflated Poisson, ZINB - Zero-inflated Binomial, PH - Poisson hurdle, NBH - Negative binomial hurdle.

Model	df	AIC	BIC
P	3	5586.637	5600.049
NB	4	3562.385	3580.268
ZIP	6	4984.583	5011.408
ZINB	7	3534.337	3565.633
PH	6	4984.728	5011.553
NBH	7	3536.046	3567.342

indicates that these two models fit significantly better by any of the remaining models. We continue by interpreting the coefficients of the ZINB model.

Hypothesis test. The ZINB model does not show a significant association of code sharing and whether or not an article is cited at all (Table 4). Moreover, the association of code sharing with the citation counts is not statistically significant either. Thus, we ought not to reject the null hypothesis that the code sharing is not associated with citation counts.

Table 4: Regression coefficients for both parts of the ZINB regression model (count model and zero model) with outcome “citation rate”, independent variable code sharing and a covariate publication year.

Part	Parameter	Estimate	SE	Z	<i>p</i>	Lower 95% CI	Upper 95% CI
Count	(Intercept)	3.147	0.115	27.258	<0.001	2.921	3.374
	Code Sharing	0.163	0.090	1.809	0.071	-0.014	0.340
	Publication Year	-0.268	0.023	-11.641	<0.001	-0.313	-0.223
	Log(theta)	0.259	0.083	3.104	0.002	—	—
Zero	(Intercept)	-12.228	2.721	-4.493	<0.001	-17.562	-6.894
	Code Sharing	-0.022	0.442	-0.049	0.961	-0.889	0.845
	Publication Year	1.463	0.346	4.224	<0.001	0.784	2.143

To ease the interpretation of the coefficients and provide a measure of effect size, rate ratios were calculated. As count regression models violate linear regression assumptions, rate ratios are used as the standard effect size measure of count models (Coxe, 2018), calculated as e^β (Atkins, Baldwin, Zheng, Gallop, & Neighbors, 2013). Just like odds ratios, rate ratios are inversely proportional around one (Atkins et al., 2013). The rate ratio for the intercept indicates the baseline, that is, the estimated citation rate for articles that do not share code (when all other predictors are equal to 0) (Atkins et al., 2013). For rate ratios for the different predictors, the distance above or below 1 is interpreted as the percentage increase or decrease in the outcome for a one unit increase in the predictor (Atkins et al., 2013, p.6). In other words: for one unit increase in the predictor 1, we expect the predicted outcome to be multiplied by e^{β_1} (Coxe, 2018).

The maximum likelihood estimate of the rate ratio indicates that a adoption of code sharing results in a multiplication of the outcome variable citation counts by 1.177; given that an article does not come from

Figure 5: The citation counts plotted as a function of publication year and code sharing (the points are jittered along the x-axis to make the individual articles more visible). The lines depict the predicted values of citation counts from the zero-inflated Negative binomial model.

a population that does not receive any citations, code sharing is, on average, associated with an increase of citation counts by about 18 %. Figure 5 displays the observed citation counts and the expected values of the citation counts.

Multiverse analysis. Given that our analysis contains several statistical choices (e.g., excluding outliers, model selection), we conducted a multiverse analysis to explore the robustness of our conclusion. The analysis can be found at <https://osf.io/kja9y/>.

In this analysis, we fitted the six models on the complete data set, calculated the Pearson’s residuals for each model, and refitted each model while excluding data points with the Pearson’s residuals larger than 2, 3, 5, and 10. This results in 6 (models) $\times$ 6 (outlier exclusion: none, default `boxplot`, Pearson $>$ 2, 3, 5, and 10). We extracted the regression coefficients for code sharing and their associated p-values for all 36 models. Table 5 summarises the results. The sign is consistently positive across all models, but the size of the estimate somewhat varies (0.09–0.287, or increase of the number of citation counts of about 9 % to 33 % when code is shared, and when taking into account the probability that the citation count will be zero, about 10 % – 32 %). However, the coefficients are consistently statistically significant only for the Poisson based models. Given that the data manifest overdispersion, Poisson regression probably underestimates the standard errors of the coefficients – thus, we believe that rejecting the null hypothesis is indeed unwarranted.

Table 5: Regression coefficients (p -values in parentheses) of code sharing for different analytic choices.

Model		Outlier removal					
Counts	Zeros	None	Default	Z2	Z3	Z5	Z10
Negative Binomial	Hurdle	0.098 (0.489)	0.155 (0.111)	0.144 (0.18)	0.096 (0.414)	0.287 (0.019)	0.21 (0.122)
	Zero-inflated	0.102 (0.356)	0.163 (0.071)	0.157 (0.099)	0.095 (0.341)	0.232 (0.023)	0.237 (0.027)
	None	0.114 (0.291)	0.187 (0.041)	0.141 (0.13)	0.107 (0.267)	0.24 (0.016)	0.277 (0.008)
Poisson	Hurdle	0.097 (<0.001)	0.09 (0.007)	0.155 (<0.001)	0.138 (<0.001)	0.223 (<0.001)	0.095 (<0.001)
	Zero-inflated	0.097 (<0.001)	0.09 (0.007)	0.154 (<0.001)	0.127 (<0.001)	0.223 (<0.001)	0.095 (<0.001)
	None	0.117 (<0.001)	0.122 (<0.001)	0.131 (0.002)	0.121 (<0.001)	0.149 (<0.001)	0.17 (<0.001)

Table 6: Comparison of the six models: P - Poisson, NB - Negative binomial, ZIP - Zero-inflated Poisson, ZINB - Zero-inflated Binomial, PH - Poisson hurdle, NBH - Negative binomial hurdle.

Model	df	AIC	BIC
P	3	1972.005	1985.313
NB	4	1970.116	1987.861
ZIP	6	1975.769	2002.386
ZINB	7	1974.346	2005.399
PH	6	1908.764	1935.381
NBH	7	1812.926	1843.979

3.4 Association of code sharing with Altmetric Attention Score

The analysis with AAS as the dependent variable follows the same plan as the analysis of citation counts, with the only exception of using AAS instead of citations as the dependent variable.

Model selection. Table 6 shows that the negative binomial hurdle model has the lowest AIC and BIC. This indicates that relative to other models, the negative binomial hurdle model performs the best. The Vuong test confirmed that the NBH model performs significantly better than all alternatives. The rootograms for the NBH model showed a superior fit of the NBH model compared to other models. Thus, we continue interpreting the coefficients of the NBH model.

Hypothesis test. The results of the NBH model with outcome AAS, the predictor code sharing, and the control variable publication year, show a significant association between code sharing in the count part of the model, $\beta = 0.380$, $p = .023$ (Table 7); code sharing is positively related to the number of mentions in alternative sources.

The implied rate ratios indicate that code sharing is associated with a multiplication of the outcome variable AAS by 1.462; given a particular article does not have zero AAS, having shared code is associated, on average, with increase of AAS by 46 %. Taking into account the probability of article having a zero AAS, sharing code is associated with an expected increase of AAS by about 17 %. Figure 6 displays the observed AAS plotted against the model expectations (i.e., expected values of the predictions from the model).

Figure 6: The AAS plotted as a function of publication year and code sharing (the points are jittered along the x-axis to make the individual articles more visible). The lines depict the predicted values of AAS from the Negative binomial hurdle model.

Multiverse analysis. We also conducted the multiverse analysis, which was done according to the same criteria as with Citation Counts. The analysis is available at <https://osf.io/xc4d/>.

Table 8 summarises the results. The sign is consistently positive across all models, but the size of the estimate varies substantially (0.159–0.647, or increase of the number of citation counts of about 17 % to 91 % when code is shared; when taking into account the probability that AAS is zero, the range is between 17 % – 85 %). Furthermore, the coefficients are consistently statistically significant, regardless of the model choice or exclusion criteria. Thus, we consider the conclusion that the null hypothesis ought to be rejected as relatively robust against these modelling choices, although the effect size is considerably unstable.

Table 7: Regression coefficients for both parts of the NBH regression model (count model and zero hurdle model) with outcome “Altmetric Attention Score”, independent variable code sharing and covariate publication year.

Part	Parameter	Estimate	SE	Z	<i>p</i>	Lower 95% CI %	Upper 95% CI %
Count	(Intercept)	-1.622	0.808	-2.007	0.045	-3.206	-0.038
	Code Sharing	0.380	0.167	2.267	0.023	0.051	0.708
	Publication Year	0.061	0.036	1.718	0.086	-0.009	0.131
	Log(theta)	-1.649	0.976	-1.691	0.091	-	-
Zero	(Intercept)	0.226	0.269	0.842	0.400	-0.300	0.753
	Code Sharing	0.028	0.273	0.103	0.918	-0.507	0.563
	Publication Year	0.396	0.062	6.430	<0.001	0.275	0.516

Table 8: Regression coefficients (p -values in parentheses) of AAS for different analytic choices.

Model		Outlier removal					
Counts	Zeros	None	Default	Z2	Z3	Z5	Z10
Negative binomial	Hurdle	0.516 (0.009)	0.38 (0.023)	0.527 (0.003)	0.647 (<0.001)	0.557 (0.002)	0.407 (0.025)
	Zero-inflated	0.272 (0.016)	0.159 (0.022)	0.285 (<0.001)	0.324 (<0.001)	0.352 (<0.001)	0.189 (0.048)
	None	0.272 (0.011)	0.166 (0.016)	0.285 (<0.001)	0.324 (<0.001)	0.352 (<0.001)	0.189 (0.045)
Poisson	Hurdle	0.354 (<0.001)	0.289 (<0.001)	0.454 (<0.001)	0.4 (<0.001)	0.324 (<0.001)	0.396 (<0.001)
	Zero-inflated	0.354 (<0.001)	0.159 (0.016)	0.342 (<0.001)	0.305 (<0.001)	0.291 (<0.001)	0.386 (<0.001)
	None	0.343 (<0.001)	0.166 (0.011)	0.617 (<0.001)	0.296 (<0.001)	0.287 (<0.001)	0.305 (<0.001)

4 Conclusions and Discussion

The current study was conducted with a twofold purpose, namely (1) to gain insight into the current status of code sharing practices in the field of psychological methods and statistics, and (2) to investigate the association of code sharing with both a conventional research metric (citation counts) and an alternative research metric (Altmetric Attention Score) within the same academic field.

We have found that sharing scientific computer code is not (yet) extensively practiced within the field of psychological methods and statistics. Out of the number of articles that used code, about 66 % did not contain any form of code sharing. Nevertheless, the findings of this study also suggest a hopeful prospect as the overall sharing of scientific computer code has increased between 2010 and 2017. Specifically, Journal of Psychological Methods exhibited a rapid decrease of the number of published articles that did not share code.

When code sharing was practiced, code was often incomplete (≈ 53 %) and/or annotated in an improper or fragmentary manner (≈ 70 % of articles that shared code provided no or deficient annotation or documentation). This likely hinders reproducibility of the reported results, and certainly limits the potential to understand and reuse the code by other researchers (Epskamp, 2019). Lastly, code or software that was reused by other researchers was not properly cited in about half of the articles. We believe that requiring proper acknowledgment is crucial, as writing code is becoming an integral (and ever more time-consuming) part of many scientists' work. Citing others' code is not only implied by ethics of scientific publishing, but could play an important role as a motivator for researchers to share their own code openly.

Concerning the association of code sharing with the conventional research citation rate and the alternative research metric Altmetric Attention Score, the results of this study did not find an association between code sharing and citation rate, but revealed a positive association between code sharing and Altmetric Attention Score. These findings are supported by previous research into open science practices which found similar increases in the number of mentions in alternative sources (Adie, 2014; McKiernan et al., 2016; Wang, Liu, Mao, & Fang, 2015).

4.1 Limitations

This study has several limitations. The first of these limitations is the potential presence of rater bias. All articles in the pilot study and main study were coded by a single rater, despite all our effort to standardize the procedure (Hoyt, 2000).

A second limitation of this study is the failure to control for all variables that were shown to correlate with research metrics in previous research (Piwowar & Vision, 2013). According to Piwowar and Vision (2013), who studied the association of data sharing with citation rate, the most important correlating variables are “date of publication, journal that published the study, journal impact factor, journal citation half-life, number of articles published by the journal, journal open access policy, journal status as a core clinical journal by MEDLINE, number of authors of the study, country of the corresponding author, citation score of the institution of the corresponding author, publishing experience of the first and last author, and subject of the study itself”. The present study however includes only publication year as a covariate. As a consequence, variance that now seems to be explained by the variable code sharing may actually be explained by one of the covariates as proposed by Piwowar et al. (2007); Piwowar and Vision (2013).

A third limitation of this study is that a weighted sum score of the alternative research metrics was used. As a consequence, this study does not provide insight in the association between code sharing and the individual alternative metrics.

Further, the nature of this study is purely observational, yielding little insights into the causal mechanisms. For example, the Open Science movement gained on momentum partly through social media (e.g., through public Facebook groups or active Twitter community). It is possible that people active in these communities have stronger inclinations towards adopting open science practices, as well as discussing own or each others work (in other words, increased alternative metrics may merely reflect that people who already endorse and practice open science tend to discuss open science on social media).

4.2 Recommendations

Despite displaying a positive relationship with alternative research metrics, it appears that code sharing is unfortunately not (yet) a common practice within the field of psychological methods and statistics. The findings of this study suggest a number of recommendations with regard to open science. In particular, they demonstrate the need for more active promotion of code sharing by the open science movement. By emphasizing the personal benefits for researchers, (i.e. increased alternative research metrics), researchers can hopefully be motivated to share code in a more frequent and comprehensive manner.

However, should this prove to be insufficient incentivising, code sharing could also be achieved by al-

ternative (or even more coercive) means. Here, an important role is reserved for funders, publishers, and institutions. Funders and publishers could use their policies on open science practices to actively promote code sharing. Previous research on the promotion of the open science practice data sharing has shown that, if implemented, researchers are willing to comply to mandatory data sharing policies. This suggests that such policies will not result in authors seeking alternative funders or publishers who do not mandate such sharing (Houtkoop et al., 2018). Whereas several funders and publishers require their authors to provide a data management plan (Houtkoop et al., 2018), the same policy could be implemented with regard to code sharing.

In addition to such policies, scientific institutions could also provide educational materials and training on the topic of code sharing. Further, ensuring that a code is useable and reproducible (in the future) is not a trivial task, which requires additional knowledge (Epskamp, 2019). This resonates with previous research on open science practices, which has found that insufficient knowledge on data sharing practices often forms an important barrier to sharing primary research data (Houtkoop et al., 2018). This finding can also be extended to code sharing, where the provision of educational material (for example on the availability of code repositories) could significantly contribute to the use open science practices.

The beneficial association between code sharing and alternative research metrics will hopefully provide a new incentive for researchers in the field of psychological methods and statistics to engage more extensively in the practice of code sharing. Nevertheless, while such incentives may prove an invaluable motivational tool for researchers to practice open science, it is vital that they do not compose the main reason for its existence. Indeed, as Piwowar and Vision (2013) note: “it is important to remember that the primary rationale for making research data available has nothing to do with evaluation metrics or citation benefits: giving a full account of experimental process and findings is a tenet of science, and publicly-funded science is a public resource (Smith, 2006)”. What is to be taken from this statement is that open science is not ultimately about achieving the highest amount of exposure or attention. Rather, it is about holding oneself to the highest academic standards, while applying the principles of openness to the whole research cycle and all types of primary research outputs.

Contributions

Ingmar Visser provided the conceptual idea of this article. Bobby-Lee Houtkoop and Ingmar Visser designed the study. Bobby-Lee Houtkoop conducted the literature search and data collection, descriptive analyses, hypothesis testing, and drafted the article. Šimon Kucharský reviewed the code, analyses and materials, conducted the multiverse analysis, and finalized the article and online supplementary materials. All authors

contributed to the final editing of the manuscript.

Acknowledgments

Šimon Kucharský was supported by the NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) grant no. 406.10.559. We would like to thank Altmetric for cooperation (answering our questions regarding their system, and enabling us to share their data openly).

Open Practices

The data and analysis code have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/4keaq/>. The design and initial analysis plan were preregistered (<https://osf.io/pc7ht/>). Our transparency report discloses the ways in which we deviated from the original plan <https://osf.io/vjme7/>. All conducted analyses are reported either in the manuscript, or enclosed in the online project repository.

References

- Adie, E. (2014). Taking the alternative mainstream. *Profesional De La Informacion*, 23(4), 349–351.
- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199–213). Springer.
- Altmetric. (n.d.-a). Sources of Attention. Retrieved July 06, 2017, from <https://www.altmetric.com/about-our-data/our-sources/news/>
- Altmetric. (n.d.-b). How it works. Retrieved July 25, 2017, from <https://www.altmetric.com/about-our-data/how-it-works/>
- Altmetric. (2016). How is the Altmetric Attention Score calculated? Retrieved July 07, 2017, from <https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-attention-score-calculated->
- Association for Psychological Science. (2016). Science Develops New Data - and Materials - Sharing Requirements. Retrieved June 02, 2017, from <https://www.psychologicalscience.org/publications/observer/obsonline/science-develops-new-data-and-materials-sharing-requirements.html>

- Atkins, D. C., Baldwin, S. A., Zheng, C., Gallop, R. J., & Neighbors, C. (2013). A tutorial on count regression and zero-altered count models for longitudinal substance use data. *Psychology of Addictive Behaviors*, 27(1), 166.
- Auguie, B. (2017). gridextra: Miscellaneous functions for "grid" graphics [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=gridExtra> (R package version 2.3)
- Barnes, N. (2010). Publish your computer code: it is good enough. *Nature News*, 467(7317), 753.
- Blohowiak, B. B., Cohoon, J., de Wit, L., Eich, E., Farach, F. J., Hasselman, F., ... Nosek, B. A. (2013). Badges to acknowledge open practices. Retrieved from <https://osf.io/tvyxz/>
- Challet, D., & Du, Y. L. (2003). Closed source versus open source in a model of software bug dynamics. *arXiv preprint cond-mat/0306511*.
- Chawla, S., Dalmeed. (2016). The unsung heroes of scientific software. *Nature News*, 529(7584), 115.
- Coxe, S. (2018). Effect size measures for nonlinear count regression models. In *Conference in statistical practice, american statistical association*.
- Epskamp, S. (2019). Reproducibility and replicability in a fast-paced methodological world. *Advances in Methods and Practices in Psychological Science*, 0(0), 2515245919847421. doi: 10.1177/2515245919847421
- Friendly, M. (2017). vcdextra: 'vcd' extensions and additions [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=vcdExtra> (R package version 0.7-1)
- Henry, L., & Wickham, H. (2019). rlang: Functions for base types and core r and 'tidyverse' features [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=rlang> (R package version 0.4.0)
- Houtkoop, B. L., Chambers, C., Macleod, M., Bishop, D. V., Nichols, T. E., & Wagenmakers, E.-J. (2018). Data sharing in psychology: A survey on barriers and preconditions. *Advances in Methods and Practices in Psychological Science*, 1(1), 70–85.
- Hoyt, W. T. (2000). Rater bias in psychological research: When is it a problem and what can we do about it? *Psychological methods*, 5(1), 64.
- Hu, M.-C., Pavlicova, M., & Nunes, E. V. (n.d.). Zero-Inflated and hurdle models for count data with extra zeros: examples from an HIV-risk reduction intervention trial. *The American journal of drug and alcohol abuse*, 37(5), 267–375.
- Jackman, S. (2017). pscl: Classes and methods for R developed in the political science computational laboratory [Computer software manual]. Sydney, New South Wales, Australia. Retrieved from <https://github.com/atahk/pscl/> (R package version 1.5.2)
- Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., ... others

- (2016). Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency. *PLoS biology*, 14(5), e1002456.
- Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 411(6837), 521–522.
- LeVeque, R. J. (2013). Top ten reasons to not share your code (and why you should anyway). *Siam News*, 46(3).
- McKiernan, E. C., Bourne, P., Brown, C., Buck, S., Kenall, A., Lin, J., ... Spies, J. R. (2016). How open science helps researchers succeed. *Elife*, 5. (e16800)
- Nimon, K., Oswald, F., & Roberts., J. K. (2013). yhat: Interpreting regression effects [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=yhat> (R package version 2.0-0)
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... others (2015). Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science*.
- Opensource.com. (n.d.). What is Open Science? Retrieved May 19, 2017, from <https://opensource.com/resources/open-science>
- Piwowar, H. A., Day, R. S., & Fridsma, D. B. (2007). Sharing detailed research data is associated with increased citation rate. *PloS ONE*, 2(3), e308.
- Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175.
- R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Robinson, D., & Hayes, A. (2019). broom: Convert statistical analysis objects into tidy tibbles [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=broom> (R package version 0.5.2)
- RStudio Team. (2018). Rstudio: Integrated development environment for R [Computer software manual]. Boston, MA. Retrieved from <http://www.rstudio.com/>
- Shamir, L., Wallin, J. F., Allen, A., Berriman, B., Teuben, P., Nemiroff, R. J., ... DuPrie, K. (2013). Practices in source code sharing in astrophysics. *Astronomy and Computing*, 1, 54–58.
- Smith, R. (2006). Commentary: The power of the unrelenting impact factor—is it a force for good or harm? *International Journal of Epidemiology*, 35(5), 1129–1130.
- Vandewalle, P. (2012). Code Sharing Is Associated with Research Impact in Image Processing. *Computing Science & Engineering*. doi: 10.1109/MCSE.2012.63
- Veldkamp, C. L., Nuijten, M. B., Dominguez-Alvarez, L., van Assen, M. A., & Wicherts, J. M. (2014). Statistical reporting errors and collaboration on statistical analyses in psychological science. *PloS*

- ONE*, 9(12), e114876.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth ed.). New York: Springer.
Retrieved from <http://www.stats.ox.ac.uk/pub/MASS4> (ISBN 0-387-95457-0)
- Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, 307–333.
- Wang, X., Liu, C., Mao, W., & Fang, Z. (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics*, 103(2), 555–564.
- Wickham, H. (2017). tidyverse: Easily install and load the 'tidyverse' [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=tidyverse> (R package version 1.2.1)
- Xie, Y. (2014). knitr: A comprehensive tool for reproducible research in R. In V. Stodden, F. Leisch, & R. D. Peng (Eds.), *Implementing reproducible computational research*. Chapman and Hall/CRC.
Retrieved from <http://www.crcpress.com/product/isbn/9781466561595> (ISBN 978-1466561595)
- Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida: Chapman and Hall/CRC. Retrieved from <https://yihui.name/knitr/> (ISBN 978-1498716963)
- Xie, Y. (2019). knitr: A general-purpose package for dynamic report generation in R [Computer software manual]. Retrieved from <https://yihui.name/knitr/> (R package version 1.23)
- Zeileis, A., & Kleiber, C. (2018). countreg: Count data regression [Computer software manual]. Retrieved from <http://R-Forge.R-project.org/projects/countreg/> (R package version 0.2-1)
- Zeileis, A., Kleiber, C., & Jackman, S. (2008a). Regression models for count data in R. *Journal of Statistical Software*, 27(8). Retrieved from <http://www.jstatsoft.org/v27/i08/>
- Zeileis, A., Kleiber, C., & Jackman, S. (2008b). Regression models for count data in R. *Journal of Statistical Software*, 27(8), 1–25. doi: 10.18637/jss.v027.i08
- Zhu, H. (2019). kableextra: Construct complex table with 'kable' and pipe syntax [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=kableExtra> (R package version 1.0.1)