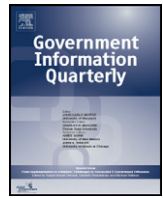




Contents lists available at ScienceDirect

# Government Information Quarterly

journal homepage: [www.elsevier.com/locate/govinf](http://www.elsevier.com/locate/govinf)



## Measuring scholarly use of government information: An altmetrics analysis of federal statistics



Tara Das

Government Information Librarian, Columbia University, Lehman Social Sciences Library, 420 W. 118th St., New York, NY 10027, USA

### ARTICLE INFO

*Article history:*

Received 4 December 2014  
 Received in revised form 13 May 2015  
 Accepted 19 May 2015  
 Available online 7 June 2015

*Keywords:*

Citation analysis  
 Government information  
 Altmetrics  
 Statistics  
 Content analysis

### ABSTRACT

**Purpose:** This paper examines how federal statistics is used in scholarly research via a new type of citation analysis that leverages the strengths of information aggregators, like Altmetric LLP, in looking for evidence of government information use beyond traditional citations/references. In this citation analysis, abstracts were examined.

**Results:** Drawing on a dataset containing articles aggregated via Altmetric Explorer, a querying interface provided by Altmetric LLP, content analysis was used to 1) determine the distribution of federal statistics incorporated in scholarly studies, and 2) qualitatively understand the particular ways in which studies incorporated federal statistics. It was found that the dominant source of federal statistics was the National Center for Health Statistics (NCHS), followed by the Census Bureau, and then the Bureau of Labor Statistics. Prevalent qualitative themes underlying the studies in this dataset included mortality and population studies, linked datasets, international studies, and critical studies (i.e. presenting alternative measures for federal statistics).

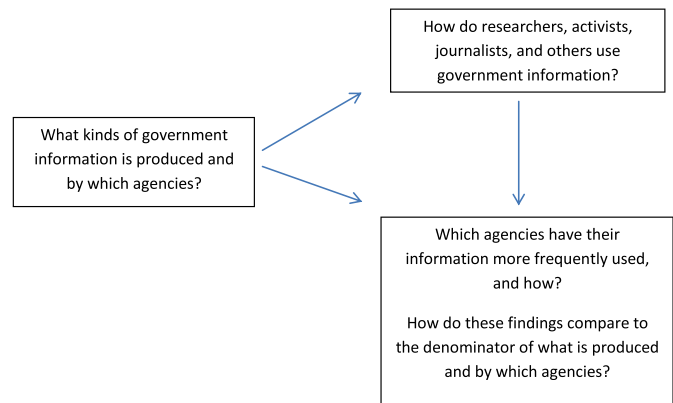
**Conclusions:** When querying studies referencing one or more of the principal US statistical agencies in Altmetric Explorer, almost all studies in the final dataset cited these agencies because they had cited federal statistics. This finding need not have been the case however. A separate study on the use of federal statistics in scholarly research will compare altmetrics to traditional citation analysis.

Preliminary results from Google Scholar, using traditional citations, found non-dataset publications to be the most frequently cited titles from NCHS and the Census Bureau.

© 2015 Elsevier Inc. All rights reserved.

### 1. Introduction

The research agenda underlying this study is to measure and analyze the production, dissemination, and use of government information in the United States. Given the wealth and variety of information that the US government produces, several studies are planned to address this agenda in order to understand the following: How much government information is produced? What exactly is produced? Which particular agencies or bureaus are most prolific in producing government information? How is government information used by others (e.g. researchers, activists, journalists)? Which agencies or bureaus have their information more frequently used by others? These questions are complex and can be measured in different ways, focusing on specific types and users of government information. The diagram below visualizes the relationships among the primary questions.



Within the research agenda, this article focuses on how researchers use federal statistics, produced by principal statistical agencies, in their studies; and which statistical agencies have their information more frequently used. It will investigate how federal statistics are used in scholarly research. It will use a new type of citation analysis that leverages

E-mail address: [tdas@columbia.edu](mailto:tdas@columbia.edu).

the strengths of altmetrics, in looking for evidence of government information use beyond traditional citations and references. It will also apply content analysis to understand the ways in which federal statistics are incorporated into scholarly studies, i.e. how this government information was used, which is a new contribution to the metrics literature.

## 2. Background

Federal statistics is a major category of US government information. Overall, US government information “constitute[s] a great library covering almost every field of human knowledge and endeavor” (Schmeckebier, Eastin, & Brookings Institution, 1969, p. 1). In particular, federal statistics “present statistical pictures of conditions and afford bases for measuring social and economic change” (Ibid.), such as unemployment rates, high school graduation rates, leading causes of death, and total population counts. These kinds of statistics provide context for identifying problems, allocating public resources, and assessing program effectiveness. Importantly, many federal statistics that the US government produces cannot be replicated by nonprofit organizations and/or universities, as individuals and organizations can be required to report particular events (e.g. births or deaths) or complete questionnaires (e.g. Decennial Census) under legal mandate. As such, federal statistics are highly utilized in scholarly research as raw data for analysis.

Highlighting the importance of federal statistics to scholarly research, journalism, and public policy, a special issue of the *Annals of the American Academy of Political and Social Science* was dedicated to “The Federal Statistical System: Its Vulnerability Matters More than You Think” in September 2010. In the introductory article, Prewitt (2010) states,

In particular - and the justification for this volume of *The Annals* - note that the empirical social sciences, from which we get much of the social knowledge relevant to public policies, would not have reached current levels of maturity in the absence of public statistics...federal statistics are indispensable to the scientific investigation of a significant number of social processes, structures, and behaviors. These investigations in turn contribute social knowledge found to be useful in policy design, implementation, evaluation, and adjustment and in public understanding of how well things are working (pp. 7, 14).

Prewitt goes on to note that a JSTOR search of leading social science journals for articles from 2008–09 demonstrated that more than half of them used statistics from five US government agencies: the Bureau of Economic Analysis (BEA), the Bureau of Labor Statistics (BLS), the U.S. Census Bureau, the National Center for Health Statistics (NCHS), and the National Oceanic and Atmospheric Administration (NOAA).

Unfortunately, Prewitt does not specify the methods used to collect these articles and analyze their citations. Notably, this exclusion is reflected in other studies that measured scholarly use of government information. There have been a few citation studies – counting and analyzing citations in a specific journal or group of journals – conducted on government information. For instance, Goehlert (1979) analyzed the use of US and international government documents in articles published in the journal *International Organization* from 1972–1976, and found the most frequently cited documents were issued by the State Department and Congress. Brill (1990) also conducted a citation analysis within international relations journals for 1964, 1974, and 1984, which indicated that 46% of all documents cited were from the US government. While detail is provided on the selection process for journals and disciplines in these citation studies (also see Hernon & Shepherd, 1983; Hogenboom, 2002), sufficient detail on how US government documents were identified in the citations is lacking.

As another weakness of traditional citation analysis, it is evident that many scholarly articles do not cite government information in a uniform

fashion. They may only cite the government source within the body of the article, and not in the references as a formal citation. When formal citations exist, the authoring agency cited for a particular information product is also not consistent. For instance, Centers for Disease Control and Prevention (CDC) or National Center for Health Statistics (a center within the CDC) may be cited as the author for annual vital statistics reports. The lack of consistency around dataset citation practices compounds the difficulty of using traditional citation analysis to measure use of federal statistics.

Moreover, inconsistencies in citing government information make it even more critical for authors of citation studies to document how they identified articles for citation analysis. In addition, how scholarly articles incorporate government information, and which federal agencies are more frequently cited, are research questions that are necessary to analyze in order to better understand how government information is used. Citation studies that have focused on government information, infrequent in themselves, have primarily focused on raw counts and percentages of citations to US government information over all citations. In one exception, Hernon and Shepherd (1983) did find that the majority of US government information, in social science citations, was issued by the Census Bureau, Congress, Federal Bureau of Investigation, National Center for Health Statistics, and Office of the Federal Register. While they point out that social scientists do refer to government information for federal statistics, their study did not discern how many of the citations were to datasets.

Given that federal statistics are used in scholarly research, and are in the form of datasets, traditional citation analysis may fail in capturing their usage. Datasets are infrequently cited in bibliographies and reference lists (though there are organizations attempting to build standards for data citation, like DataCite and Inter-university Consortium for Political and Social Research). “Traditional metrics have generally dealt with journals or articles and not measured other significant research output like blog posts, slideshows, datasets, and other important scholarly dialog” (Galligan & Dyas-Correia, 2013, p. 56).

Fortunately, alternative metrics (or altmetrics) have been developed to help address these issues. They employ different ways of capturing citations, since so much of scholarly products are now disseminated in electronic journals and other Internet publishing platforms. The emphasis thus far in altmetrics studies has been on ways in which altmetrics can capture online scholarly communication and impact of scholarly articles in social media (e.g. via number of tweets and retweets on Twitter, number of Facebook or blog posts). “In the online environment, we track almost every movement: clicks, page views, and interactions as well as the way we share things with others. A great opportunity lies in the capture of the resulting data trail and building meaningful layers of insights onto it” (Galligan & Dyas-Correia, 2013, p. 58).

In addition, libraries can use this data for collection development purposes, particularly for electronic resources available online like datasets. For these resources, circulation statistics, or even number of hits, do not provide much help in understanding which ones should be highlighted in research guides and/or prioritized for digital preservation.

Using publicly available data gathered from APIs from broad and sector-specific networks is something that cannot be ignored. Altmetrics present an integrated view of how unit of content or one researcher has moved across the digital landscape in a series of actions or digital conversations (Galligan & Dyas-Correia, 2013, p. 58).

As Priem, Piwowar, and Hemminger (2012b), note “In growing numbers, scholars are integrating social media tools like blogs, Twitter, and Mendeley into their professional communications. The online, public nature of these tools exposes and reifies scholarly processes once hidden and ephemeral.” Moreover, citations are slower to accumulate than

Download English Version:

<https://daneshyari.com/en/article/1024340>

Download Persian Version:

<https://daneshyari.com/article/1024340>

[Daneshyari.com](https://daneshyari.com)