From
Sighting Data
to
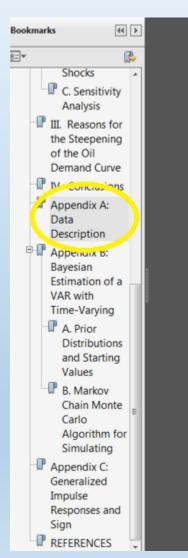# Citing Data

Alison Raab Labonte

Data Librarian, Research Library

Board of Governors of the Federal Reserve System

7 October 2016

Beyond the Numbers Conference, Federal Reserve Bank of St. Louis

Data Citing… or Data Sighting?

Citations in Academic Research

Research Citations: "Coin of the academic realm" (Kratz (2015))

- H-Index; Web of Science Citation Index; Google Scholar

- NISO, Journal Article Markup Suite (JATS) ANSI/NISO Z39.96-2015

Does data count in citation analysis?

- The good news:
  - Research that cites data, gets cited more (Piwowar, Day & Fridsma (2007))

- The bad news: Low compliance with recommendations
  - Social science papers cite data 17% of the time (Mooney (2011))
  - My unscientific survey of AEA journals – 23% list data in references

Data as the "Dark Matter" of the Scholarly Ecosystem

"Data has be... data citation aims to make...
Altman (2016)

Data citation... has played in ensuring **long**... **permanent access** and e...
Rans, et al. (2013)

"On the surfa... he **process of making data**... nsequence of this and othe... **w to develop**."
Ball and Duke (201...

Universe content
visible matter 5%
dark matter 27%
dark energy 68%

AEA recommendations for data citation

- American Economics Association (AEA)
  - https://www.aeaweb.org/journals/aer/submissions/accepted-articles/styleguide
  - Sample references: https://www.aeaweb.org/journals/policies/sample-references

Key players in data citation standards

- DataCite – https://www.datacite.org/
  - Metadata schema (updated Sept 2016): http://schema.datacite.org/
  - Useful for: a data citation schema you can implement.
  - Its schema is used by ICPSR, Dataverse, etc.
  - Handy data citation formatter! https://www.datacite.org/citation.html

- CoDATA http://www.codata.org/ and Force11 https://www.force11.org/about
  - Joint Declaration of Data Citation Principles: https://www.force11.org/group/joint-declaration-data-citation-principles-final
  - Data Citation Primer: http://force11.github.io/data-citation-primer/authors/
  - Useful for: theory and major principles behind data citation.

- IASSIST: http://www.iassistdata.org/topic/other-topics/data-citation
  - Handy quick guide to data citation: http://iassistdata.org/sites/default/files/quick_guide_to_data_citation_high-res_printer-ready.pdf

- ICPSR: https://www.icpsr.umich.edu/icpsrweb/ICPSR/curation/citations.jsp

- California Digital Library: https://datapub.cdlib.org/datacitation/

Elements of Data Citation

CORE	
Author/creator	the person(s), organization, agency responsible for creating the dataset
Year of Publication	the year the dataset was made available (not the date coverage of the data)
Title or description	complete title (if no title exists, create a brief description of the data)
Publisher	entity (organization, database, archive, journal) that hosts/provides/licenses the data
DOI/URI	the unique identifier if the data set is online (Not so easy for some data sources, such as vendor data)

ADDITIONAL	
Location/Availability	
Version/Edition	
Access date	Crucial for reproducing analysis of continuously updated dynamic data. Also useful for vendor/commercial data, as vendor names change.

What can you do:
Adapt a Data Citation Standard

Example at the Board – modified DataCite, Chicago Manual of Style (16th Ed.), and local editing practices.

What can you do:
Adapt a Data Citation Standard

Creator [(Year published)]. *Title of Data Set.* [Publisher] [, accessed month/year] [, DOI or URL].

For the confidential PSID data:

> Panel Study of Income Dynamics (2014). Restricted use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI., accessed May 2015. http://psidonline.isr.umich.edu/Guide/FAQ.aspx?ID=1040.

For the CoreLogic data:

> Corelogic, Inc. (2015). Real estate data, 2000-2015. Accessed May 2016. corelogic.com

What can you do:
Make Data Citation Easy.

Make data citation as easy as possible

- Provide recommended citations

- Export functions to EndNote, RIS, etc. (ICPSR, publishers)

- Citation formatter (DataCite/CrossRef)

- LaTeX code for easy integration into research and work

What can you do: Normalize data citation (and data deposit!)

At your organization
- Educate about the value of citation
- Provide tools to make data citation happen
- Example at the Fed: Data review team at the Board

In scholarly communication and publishing
- NISO, Journal Article Markup Suite (JATS), version 1.1, ANSI/NISO Z39.96-2015) (Standardized Markup for Journal Articles) http://jats.niso.org/1.1/
  - Lapeyre, Deborah Aleyne (2015), "Citing Data in Journal Articles Using JATS," FORCE11.
- If possible, get involved in the publishing of work products at your workplace
  - Example at the Fed: proposed transition of FEDS Working Papers from html to a repository with a linked data collection

Board Research Library

http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html

The Federal Reserve Board

Finance and Economics Discussion Series

FEDS Working Papers

The U.S. Treasury Yield Curve: 1961 to th...

Refet S. Gurkaynak, Brian Sack, and Jonatha...
2006-28

Abstract: The discount function, which determines the value of all ...
is the most basic building block of finance and is usually inferred fr...
curve. It is therefore surprising that researchers and practitioners do ...
them a long history of high-frequency yield curve estimates. This pa...
making public the Treasury yield curve estimates of the Federal Res...
frequency from 1961 to the present. We use a well-known and simp...
is shown to fit the data very well. The resulting estimates can be use...
forward rates for any horizon. We hope that the data, which are post...
http://www.federalreserve.gov/pubs/feds/2006 and which will be up...
provide a benchmark yield curve that will be useful to applied econo...

Keywords: Yield curve, forward rates, on-the-run premium, treasu...

Full paper (2532 KB PDF) | Full paper (Screen Reader Version) | Data - Excel file (30 MB XLS) | Data - Screen reader | Data - XML (sdmx/zip)

Home | Economic Research and Data | FEDS | List of 2006 FEDS papers | Accessibility | Contact us
Last update: December 22, 2011

OneBoard Research
Increasing access to Board research and data

Home | Browse All | Board Subject Taxonomy | Board Expert Finder | JEL codes | Board Research Library home | Ask-a-Librarian

Search [   ] Search Advanced Search

Home » FEDS Working Papers » The U.S. treasury yield curve : 1961 to the present / Refet S. Gürkaynak, Brian Sack, and Jonathan...

Reference URL   Add tags   Comment   Save to favorites

The U.S. treasury yield curve : 1961 to the present / Refet S. Gürkaynak, Brian Sack, and Jonathan H. Wright.

View Description   Home » Data & Addenda » Data from Gürkaynak, Sack and Wright 2006-28 FEDS paper (Excel file)

PDF | Text

160.pdf
Link will provide options to open or save d...
File Format: Adobe Reader

Reference URL   Add tags   Comment   Save to favorites

Data from Gürkaynak, Sack and Wright 2006-28 FEDS paper (Excel file)
View Description

3.xls [97.83 MB]
Link will provide options to open or save document.
File Format: Microsoft Excel - xls

Microsoft Excel

Object Description

Author	Gürkaynak, Refet S. / Sack, Brian / Wright, Jonathan H.
Title	The U.S. treasury yield curve : 196...
Date	2006
Series	Finance and Economics Discussion
Board Taxonomy	Yield curve / U.S. Treasuries
Author's keywords	Yield curve ; Forward rates ; On-th...
Pagination	31 p.
Abstract	The discount function, which determin... that researchers and practitioners do r... Federal Reserve Board at a daily frequ... compute yields or forward rates for any horizon. We hope that the data, which are posted on the website http://www.federalreserve.gov/pubs/feds/2006 and which will be updated pe... will provide a benchmark yield curve that will be use... to applied economists.
Publisher	Board of Governors of the Federal Reserve Syste... (U.S.)
Cite as	Gurkaynak, Refet S., Brian Sack, and Jonathan H. ...ght (2006). "The U.S. Treasury Yield Curve: 1961 to the Present," Finance and Economics Discussion Series 2006-28. Board of Governors of... Federal Reserve System (U.S.).
Supporting data	(Excel): http://cdm16798.contentdm.oclc.org/cdm/ref/collection/p16798coll6/id/2... (XML): http://cdm16798.contentdm.oclc.org/cdm/ref/collection/p16798coll6/id/3...
Later versions	(2007, Journal of Monetary Economics) http://cdm16798.contentdm.oclc.org/cdm/ref/collection/p16798coll7/id/2019

Description

Title	Data from Gürkaynak, Sack and Wright 2006-28 FEDS paper (Excel file)
Author	Gürkaynak, Refet S. / Sack, Brian / Wright, Jonathan H.
Date	2006
Board Taxonomy	Yield curve / U.S. Treasuries
Series	Finance and economics discussion series ; 2006-28
Publisher	Board of Governors of the Federal Reserve System (U.S.)
Relation	The U.S. treasury yield curve : 1961 to the present / Refet S. Gürkaynak, Brian Sack, and Jonathan H. Wright: http://cdm16798.contentdm.oclc.org/cdm/ref/collection/p16798coll4/id/201

Incentivize Data Citation.

"Academic researchers as a class are drawn to research and scholarship through an interest in puzzle-solving, but they are also substantially incented by recognition and money." (Altman 2016)

• Make data citation valuable to your users and organization

• Value may differ for: economists; research assistants; data managers; librarians; budget team.

Thanks!

Any questions or comments?

Alison Raab Labonte, alison.r.labonte@frb.gov

Bibliography:

- Altman, Micah, and Gary King (2007), "A Proposed Standard for the Scholarly Citation of Quantitative Data," *D-Lib Magazine*, (April 2007) http://www.dlib.org/dlib/march07/altman/03altman.html

- Altman, Micah (2016), "The Role of Research Funding and Policy Community in Data Citation — Rewards, Incentives, and Infrastructure." (August 24, 2016), http://informatics.mit.edu/blog/role-research-funding-and-policy-community-data-citation-%E2%80%94-rewards-incentives-and

- American Economic Association, "AEA recommendations on Citation,) https://www.aeaweb.org/journals/aer/submissions/accepted-articles/styleguide

- Ball, Alex and Monica Duke (2011), "How to Cite Datasets and Link to Publications," (July 19), http://opus.bath.ac.uk/28006/1/idccws6-ab-slides.pdf

- Buneman, Peter; Susan Davidson, and James Frew (2016), "Why Data Citation is a Computational Problem," *Communications of the ACM*, vol. 59, No. 9, pp. 50-57. http://cacm.acm.org/magazines/2016/9/206243-why-data-citation-is-a-computational-problem/fulltext

- Chavan, Vishwas, and Lyubomir Penev (2011), "The data paper: a mechanism to incentivize data publishing in biodiversity science," *BMC Bioinformatics,* 12 (supl 15): S2 (15 December 2011), https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-S15-S2

- DataCite. Data Citation Schema, last updated September 2016), http://schema.datacite.org

- Executive Office of the President (2013), Office of Science and Technology Policy, (2013), "Memorandum of the Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research," (2/22/2013), https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

- Force11 (2014), "Data Citation Synthesis Group: Joint Declaration of Data Citation Principles," Martone M. (ed.) San Diego CA: FORCE11; 2014 (https://www.force11.org/group/joint-declaration-data-citation-principles-final]

- Hole, Brian (2013), "Incentivizing Data Sharing: The Journal of Open Public Health Data," Presentation at the Workshop for Open Access for Health Scientists and Professionals, Kampala (30 August 2013),   http://www.slideshare.net/brianhole/ubiquity-press-eifl130830

- IASSIST, Data Citation Quick Guide. http://www.icpsr.umich.edu/files/ICPSR/enewsletters/iassist.html

- Kratz, John E. and Carly Strasser (2015), "Making data count," *Sci. Data* 2:150039. http://www.nature.com/articles/sdata201539

- Kunze, John, Robert Cook (2009), "Defining the Data Citation Problem in the DataNet Context." DataONE Presentation (December 2009), *openwetware.org/images/3/37/Datanet_citation_agu_jak.ppt*

- Lapeyre, Deborah Aleyne (2015), "Citing Data in Journal Articles Using JATS," FORCE11, https://www.force11.org/sites/default/files/d7/project/882/citing-data-in-jats-2015-06.pdf

Bibliography (continued)

- Michigan State University (MSU), Data Citation Guide, http://libguides.lib.msu.edu/citedata

- Mooney, H. (2011), "Citing data sources in the social sciences: do authors do it?" *Learned Publishing* 24, (2011), pp. 99–108, https://pdfs.semanticscholar.org/c52e/db2eb663e1e43f5ad03738d73923307566f6.pdf

- Mooney, H, MP Newton (2012), "The Anatomy of a Data Citation: Discovery, Reuse, and Credit," *Journal of Librarianship and Scholarly Communication,* 1(1):eP1035. http://dx.doi.org/10.7710/2162-3309.1035

- Moss, Elizabeth, Christin Cave, Jared Lyle (2015), "Sharing and citing research data: A repository's perspective," https://deepblue.lib.umich.edu/handle/2027.42/115490

- NISO, Journal Article Markup Suite (JATS), version 1.1, ANSI/NISO Z39.96-2015 (Standardized Markup for Journal Articles) http://jats.niso.org/1.1/

- Pienta AM, GC Alter GC, and JA Lyle (2010), "The enduring value of social science research: the use and reuse of primary research data," https://deepblue.lib.umich.edu/handle/2027.42/78307

- Piwowar, Heather A., Roger S. Day, Douglas B. Fridsman (2007), "Sharing Detailed Research Data Is Associated with Increased Citation Rate," *PLOS* (March 21, 2007), http://dx.doi.org/10.1371/journal.pone.0000308

- Rans, Jonathan, Michael Day, Monica Duke, and Alex Ball (2013), "Enabling the citation of datasets generated through public health research," The Wellcome Trust, https://wellcome.ac.uk/sites/default/files/enabling-citation-of-datasets-through-public-health-research-phrdf-2013.PDF.

- Robinson-Garcia, Nicholas, Evaristo Jiménez-Contreras, & Daniel Torres-Salinas (2015), "Analyzing data citation practices according to the Data Citation Index," *arXiv:*150106285,   https://arxiv.org/ftp/arxiv/papers/1501/1501.06285.pdf

- Sieber, P.JE, BE Trumbo (1995), "(Not) Giving credit where credit is due: citation of data sets," *Science and Engineering Ethics*, 1, pp. 11–20.

- Simons, Natasha (2012), "Implementing DOIs for Research Data," *D-Lib Magazine*, Vol 18, Number 5/6 (May/June 2012), http://www.dlib.org/dlib/may12/simons/05simons.html

- US OMB Circular A-110 https://www.whitehouse.gov/omb/circulars_a110

Image credits:

Slide 3, "I [Love] Footnotes" image from: HistoryGypsy (2011), "Haunted and Taunted by My Own Wretched Foot," *While we're paused* (blog) https://lanternhollowpress.com/2011/07/10/haunted-and-taunted-by-my-own-wretched-foot/ (accessed 9/28/2016).

Slide 4: "Dark Matter" slide: http://www.quantumdiaries.org/wp-content/uploads/2013/06/disk-dark-matter.jpg

Slide 14: "Metadata is a Love Note to the Future" Quote from Jason Scott, NYPL Labs. Image from Rebekah Cummings (2015 Research data management and sharing for social and behavioral sciences and humanities Rebekah Cummings, Research Data Management Librarian J. Willard Marriott Library September 15, 2015 http://www.slideshare.net/RebekahCummings/research-data-management-and-sharing-for-the-social-sciences-and-humanities